IN COMPARING RADIATIVE TRANSFER AND CHEMICAL TRANSPORT MODELS ON OMI NO2 RETRIEVALS

Thesis
Presented to
The Academic Faculty

By

Charles David Smeltzer

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science in Earth and Atmospheric Sciences

Georgia Institute of Technology

December 2009

IN COMPARING RADIATIVE TRANSFER AND CHEMICAL TRANSPORT MODELS ON OMI NO2 RETRIEVALS

Approved by:

Dr. Yuhang Wang, Advisor
School of Earth and Atmospheric Science
Georgia Institute of Technology

Dr. Michel Chang
School of Earth and Atmospheric Science
Georgia Institute of Technology

Dr. Greg Huey
School of Earth and Atmospheric Science
Georgia Institute of Technology

Date Approved: September 22, 2009

ACKNOWLEDGEMENTS

There are several Earth and Atmospheric Science and chemistry classes that were invaluable to my education. Professor Judy Curry, Professor Rodney Weber, Professor Michael Chang, Professor Yuhang Wang, Professor Greg Huey, Professor Andrew Stack and many others have all been wonderful sources of inspiration for my own learning and teaching! In addition to my research, my master's degree experience would have been incomplete without spending three semesters as an EAS 1600 lab TA and as an EAS 4308 data analysis TA. I hope that I have served the Earth and Atmospheric Science department well in equal exchange for my experience.

I would like recognize the Earth and Atmospheric Sciences technical support, Mark Keever and Angelica I. Remolina. Their endless dedication to our group's computing needs is relentless, prescient, and expedient, installing and optimizing the functionality of our systems and work stations non-stop. I wish to thank Victoria Van Cappellen as well for technical support in writing and the logistics of this thesis.

To everyone else who has offered words of encouragement and discouragement alike, I thank you for your constructive criticisms and your faith in the strengthening of this work. This one goes out to my family and friends, my adviser, Dr. Yuhang Wang, and his research group, my reading committee and all external assistance. All your opinions mean a lot to me. I always appreciate your ideas and your wishes to make my work more informative, educational, alluring and challenging!

Thank you.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF SYMBOLS AND ABBREVIATIONS	xiii
SUMMARY	xv
CHAPTER	
1 INTRODUCTION	1
1.1 OMI Algorithm Theoretical Basis	2
1.2 NASA vs. NRT Algorithm	7
1.3 Error Analysis of Satellite Algorithm	10
1.4 Validation of KNMI and NASA	13
1.5 NO_2 Retrieval Trends, Inversions and Variability	15
1.6 Purpose for the Proposed Research	19
2 METHODS	20
2.1 Stratosphere/Troposphere Separation	20
2.2 NO_2 Radiance Test	21
2.3 A priori Vertical Column Comparison	22
2.4 Statistical Testing on Cities in July	24
3 RESULTS	28
3.1 DAK and TOMRAD Radiance Comparison	28
3.2 Constant Profile Results	31
3.3 REAM vs. TM4 Model Differences	40

3.4 Retrieval Comparison 51

3.5 Internal Retrieval issues 65

3.6 Odd vs. Even Day Statistical Results on Retrievals for Select Cities 73

4 DISCUSSION 85

5 CONCLUSIONS 88

APPENDIX: City Statistics 93

REFERENCES 105

LIST OF TABLES

Page

Table 1.2.1: Resolution and retrieval limits of the radiative transfer look-up-table 9

Table 1.3.1: Summary of Differences between KNMI and NASA Algorithms 10

Table 1.3.2: Results Sensitivity Analysis of the OMI DOAS NO2 12

Table 1.3.3: Estimated Accuracy of the OMI DOAS NO2 Algorithm 12

Table A.1: Statistics for Atlanta 36 km for 8 days starting on July 5th 93

Table A.2: Statistics for Atlanta 36 km for 8 days starting on July 1st no lightning 94

Table A.3: Statistics for Atlanta for 8 days starting on July 5th 95

Table A.4: Statistics for Chicago for 8 days starting on July 1st 96

Table A.5: Statistics for Chicago for 8 days starting on July 7th 97

Table A.6: Statistics for New York City for 6 days starting on July 5th 98

Table A.7: Statistics for New York City for 6 days starting on July 13th 99

Table A.8: Statistics for Los Angeles for 30 days starting on July 1st 100

Table A.9: Statistics for Los Angeles for 8 days starting on July 7th 101

Table A.10: Statistics for Los Angeles 36 km for 8 days starting on July 7th 102

LIST OF FIGURES

Page

Figure 1.1.1: The major absorbers within the OMI NO_2 window 4

Figure 1.1.2: The fitting window currently used by TEMIS 5

Figure 1.1.3: 60 pixel cross track of OMI, geometry and area 7

Figure 1.2.1: KNMI and NASA algorithm flowchart 8

Figure 1.3.1: Temperature correction of NO_2 a priori retrieval 11

Figure 1.3.2: Total tropospheric air mass factor uncertainty 12

Figure 1.4.1: Brewer validation of NASA OMI retrieval 14

Figure 1.4.2: SCIMACY validation of KNMI OMI retrieval 14

Figure 1.4.3: Aircraft comparison of NASA and KNMI OMI retrieval 15

Figure 1.5.1: Retrieval vs. TM4 annual maximum NO_2 concentration 16

Figure 1.5.2: Dominant type of NO_2 source 16

Figure 1.5.3: NO2 SCIAMACHY retrievals compared to REAM simulations 18

Figure 1.5.4: NO2 REAM simulations compared to INTEX-NA mission 18

Figure 2.0.1: NO_2 Retrieval Flow Chart 21

Figure 3.1.1: Differences between TOMRAD and DAK radiative models 30

Figure 3.2.1: TM4 profile mixing ratios Profiles 1-4 32

Figure 3.2.2: TM4 profile column layers Profiles 1-4 32

Figure 3.2.3: Location of the four TM4 profiles 32

Figure 3.2.4: 1-Profile1D/Profile1T July 1st 33

Figure 3.2.5: 1-Profile2D/Profile2T July 1st 33

Figure 3.2.6: 1-Profile3D/Profile3T July 1st 33

Figure 3.2.7: 1-Profile4D/Profile4T July 1st 34

Figure 3.2.8: 1-Profile1D/Profile1T July Monthly Average 35

Figure 3.2.9: 1-Profile2D/Profile2T July Monthly Average 35

Figure 3.2.10: 1-Profile3D/Profile3T July Monthly Average 35

Figure 3.2.11: 1-Profile4D/Profile4T July Monthly Average 36

Figure 3.2.12: TM4 profile mixing ratios Profiles 5-8 36

Figure 3.2.13: TM4 profile column layers Profiles 5-8 36

Figure 3.2.14: 1-Profile5D/Profile1T July 1st 37

Figure 3.2.15: 1-Profile6D/Profile2T July 1st 37

Figure 3.2.16: 1-Profile7D/Profile3T July 1st 37

Figure 3.2.17: 1-Profile8D/Profile4T July 1st. 38

Figure 3.2.18: 1-Profile5D/Profile1T July Monthly Average 38

Figure 3.2.19: 1-Profile6D/Profile2T July Monthly Average 38

Figure 3.2.20: 1-Profile7D/Profile3T July Monthly Average 39

Figure 3.2.21: 1-Profile8D/Profile4T July Monthly Average 39

Figure 3.3.1: REAM36Ms 41

Figure 3.3.2: REAM36M8 41

Figure 3.3.3: REAM36MnoLTs 41

Figure 3.3.4: REAM36MnoLT8 42

Figure 3.3.5: TM4Ms 42

Figure 3.3.6: TM4M8 42

Figure 3.3.7: REAM36M5 43

Figure 3.3.8: REAM36M3 43

Figure 3.3.9: REAM36MnL5 43

Figure 3.3.10: REAM36M3 44

Figure 3.3.11: TM436M5 44

Figure 3.3.12: TM436M3 44

Figure 3.3.13: 1-REAM36Ms/REAM36MnLs 45

Figure 3.3.14: 1-REAM36Ms/TM436Ms 45

Figure 3.3.15: 1-REAM36M5/REAM36MnL5 45

Figure 3.3.16: 1-REAM36M5/TM436M5 46

Figure 3.3.17: 1-REAM36M3/REAM36MnL3 46

Figure 3.3.18: 1-REAM36M3/TM436M3 46

Figure 3.3.19: TM4M 48

Figure 3.3.20: REAM36M 48

Figure 3.3.21: TM436M5 49

Figure 3.3.22: REAM36MnL 49

Figure 3.3.23: 1-REAM36M/REAM70M 50

Figure 3.3.24: 1-REAM36MnL/REAM36M 50

Figure 3.3.25: 1-REAM36MnL/TM436M 50

Figure 3.3.26: 1-REAM36M5/TM436M5 51

Figure 3.4.1: Comparison of the Model Vertical and Retrieval Slant Column, GA 52

Figure 3.4.2: Comparison of the Model Vertical and Retrieval Slant Column, OK 53

Figure 3.4.3: TM4D 54

Figure 3.4.4: TM4T 54

Figure 3.4.5: REAM36D 54

Figure 3.4.6: REAM36T 55

Figure 3.4.7: REAM70D 55

Figure 3.4.8: REAM70T 55

Figure 3.4.9: 1-REAM36D/REAM70D 56

Figure 3.4.10: 1-TM436M/TM436D 57

Figure 3.4.11: 1-TM436M/TM4T 57

Figure 3.4.12: 1-TM4M/TM470D 58

Figure 3.4.13: 1-TM4M/TM470T 58

Figure 3.4.14: 1-REAM36M/REAM36D 59

Figure 3.4.15: 1-REAM36M/REAM36T 60

Figure 3.4.16: 1-REAM70M/REAM70D 60

Figure 3.4.17: 1-REAM70M/REAM70T 60

Figure 3.4.18: 1-REAM70D/TM470D 61

Figure 3.4.19: 1-REAM36D/TM436D 61

Figure 3.4.20: REAM36MnL 63

Figure 3.4.21: 1-REAM36MnL/REAM36M 63

Figure 3.4.22: REAM36DnL 63

Figure 3.4.23: REAM36TnL 64

Figure 3.4.24: 1-REAM36DnL/REAM36D 64

Figure 3.4.25: 1-REAM36TnL/REAM36T 64

Figure 3.5.1: 1-RAAodd/RAAeven 5-8 66

Figure 3.5.2: 1-RAAodd/RAAeven 1-30 66

Figure 3.5.3: 1-REAM36Do/REAM36De5-8 66

Figure 3.5.4: 1-REAM36Do/REAM36De1-30 67

Figure 3.5.5: 1-REAM36Mo/REAM36Me5-8 67

Figure 3.5.6: 1-REAM36Mo/REAM36Me1-30 67

Figure 3.5.7: RAA1 69

Figure 3.5.8: REAM36D1 69

Figure 3.5.9: RAA2 69

Figure 3.5.10: REAM36D2 70

Figure 3.5.11: RAA3 70

Figure 3.5.12: REAM36D3 70

Figure 3.5.13: RAA4 71

Figure 3.5.14: REAM36D4 71

Figure 3.5.15: RAA6 71

Figure 3.5.16: REAM36D6 72

Figure 3.5.17: RAA8 72

Figure 3.5.18: REAM36D8 72

Figure 3.6.1: Distribution plot for Atlanta 36 km domain July 1st for 30 days 76

Figure 3.6.2: Distribution plot for Atlanta 36 km domain no lightning July 1st for 30 days 76

Figure 3.6.3: Auto Covariance for Atlanta 36 km domain July 1st for 30 days 76

Figure 3.6.4: Distribution plot for Chicago starting July 1st going for 8 days 78

Figure 3.6.5: Distribution plot for Chicago starting July 7th going for 8 days 78

Figure 3.6.6: Distribution plot for New York City starting July 5th going for 6 days 80

Figure 3.6.7: Distribution plot for New York City starting July 13th going for 6 days 80

Figure 3.6.8: Distribution plot for Los Angeles starting July 1st going for 30 days 82

Figure 3.6.9: Distribution plot for Los Angeles starting July 7th going for 8 days 82

Figure 3.6.10: Distribution plot for Los Angeles 36 km domain July 7th going for 8 day 84

Figure A.1: Histogram for Atlanta 36 km starting on July 1st for 30 days 93

Figure A.2: Histogram for Atlanta 36 km starting on July 1st for 30 days, no lightning 94

Figure A.3: Histogram for Atlanta starting on July 5th going for 8 days 95

Figure A.4: Histogram for Chicago starting on July 1st going for 8 days 96

Figure A.5: Histogram for Chicago starting on July 7th going for 8 days 97

Figure A.6: Histogram for New York City starting on July 5th going for 6 days 98

Figure A.7: Histogram for New York City starting on July 13th going for 6 days 99

Figure A.8: Histogram for Los Angeles starting on July 1st going for 30 days 100

Figure A.9: Histogram for Los Angeles starting on July 7th going for 8 days 101

Figure A.10: Histogram for Los Angeles starting on July 7th going for 8 days 102

Figure A.11: Viewing Zenith Angle of OMI from July 1st to July 8th 103

Figure A.12: Map of North America 104

LIST OF SYMBOLS AND ABBREVIATIONS

NO_2	Nitrogen Dioxide
AMF	Air Mass Factor (tropospheric)
VC	Vertical Column (tropospheric)
T-test	Student T-test of means
F-test	F-test of standard deviation
NASA	National Aeronautics and Space Administration
KNMI	Royal Dutch Meteorological Institute
GT	Georgia Institute of Technology
TOMRAD	NASA Radiative Model
DAK	KNMI Radiative Model

Graphics Figure Key

36	36 km resolution
70	70 km resolution
S	Surface
8	800 hPa
5	500 hPa
3	300 hPa
D	KNMI DAK
T	NASA TOMRAD
M	Model (REAM or TM4)
nL	REAM model with no lightning NOx production

Models

REAM	Regional Chemical Transport Model (GT)
TM4	Global Chemical Transport Model (KNMI)
ECMWF	European Centre for Medium-Range Weather Forecasts
WRF	Weather Research and Forecasting Model

Satellite Geometry

RAA	Relative Azimuth Angle
VZA	Viewing Zenith Angle
SZA	Solar Zenith Angle

Satellites

OMI	Ozone Monitoring Instrument
GOME	Global Ozone Monitoring Experiment
SCIAMACHY	SCanning Imaging Absorption SpectroMeter for Atmospheric ChartographY

SUMMARY

The objective of this thesis is to evaluate the sources of the differences between the NO_2 satellite retrieval products provided by the Royal Dutch Meteorological Institute (KNMI) and the National Aeronautics and Space Administration (NASA). Ground studies have shown that although both products use the same satellite, these products yield different observations for NO_2 tropospheric columns concentrations. This study does not validate either retrieval product, but rather indentifies the main sources for the discrepancy.

There are several parameters which allow successful retrieval of NO_2 vertical columns. For this study, only the difference between the radiative models and the a priori NO_2 chemical transport models were considered relevant. All other parameters, such as cloud properties, slant columns, stratospheric serration and their assumptions, were held constant. Here, the models are referred to by their proprietor's acronym: "TOMRAD" refers to the radiative model used by NASA, "DAK" refers to the radiative model used by KNMI, "TM4" refers to the a priori chemical transport model used by KNMI, and "REAM" refers to the a priori chemical transport model maintained by the School of Earth and Atmospheric Sciences at the Georgia Institute of Technology. Mixing these parameters creates four retrievals for comparison.

Many significant differences were identified after comparing these four retrievals. First, there are viewing geometry biases between the port side and the starboard side of

the satellite retrieval for each swath. These viewing geometry biases lead to artificial periodicities in the retrievals of NO_2 tropospheric vertical columns over a specific coordinate or site, such as a city. Furthermore, there were significant differences found after using different a priori NO_2 chemical transport models. The low horizontal resolution of TM4 and the satellite retrieval/TM4 coupling effect compared to REAM leads to considerable questioning of the near real time application of the KNMI NO_2 retrieval product. Though the TM4 model performs poorly, TM4 retrievals do perform nearly as well as REAM retrievals at capturing day-to-day variability and the spatial variability of the cities used as examples here. The retrievals using TOMRAD outperformed the retrievals using DAK when compared to the high resolution, hourly REAM a priori chemical transport model. In sum, these findings should lead to better optimizations of both the KNMI and NASA retrievals, and thus make their publicly available data products more reliable and accurate for general use.

CHAPTER 1

INTRODUCTION

Understanding the behavior of nitrogen dioxide (NO_2) in the troposphere is important because of its role in producing ozone via photochemistry. Excessive amounts of ozone currently pose an air quality problem to the ambient health of many cities around the world. Chemical transport models are a key first step to estimating the production of ozone and for developing mitigation strategies to avoid the adverse health effects of ozone. These models, hence the policies they promote, rely on accurate emission inventories of NO_2 and other trace gases to make accurate predictions. Such accuracy can only be bounded by in-situ observations. Because of the regional scale of the ozone air-quality problem, the best available resource for making those observations with widespread coverage and temporal dependability are satellites.

There are several NO_2 satellite retrieval products out on the market today. These include NO_2 products from the satellites GOME, SCIAMACHY, and OMI (Boersma et al. 2007). OMI is the latest product made available and it is the focus of this work. The Royal Dutch Meteorological Institute (KNMI) and the National Aeronautics and Space Administration (NASA), have both developed NO_2 retrieval products to observe NO_2 column concentrations in the troposphere (Bucsela et al. 2008). These products are widely used in chemical transport simulations in a variety of research and public advisory arenas.

Presently, there are several gaps in the understanding of how such satellite retrievals and chemical transport models are to be coupled (Boersma et al. 2008). Models rely on satellite retrievals to be constrained; nevertheless, satellite observations rely on model a priori NO_2 vertical distributions to derive the retrieval (OMI Algorithm Theoretical Basis 2002). The satellite cannot distinguish NO_2 as a function of pressure altitude. Rather, the satellite measures NO_2 as a function of the slant column area seen by its sensor. The retrieval thus relies on an a priori vertical profile to discriminate near surface or boundary layer NO_2 from that of stratospheric NO_2 aloft.

Retrievals are further complicated by sensitivities to weather conditions affecting the Earth's radiative properties. Such conditions are cloud coverage and cloud heights or seasonal albedo extent due to snow and ice coverage (Boersma 2004). Although these effects are discussed in the introduction, they are not considered here as important as the a priori and radiative models. The KNMI meteorological parameterization: cloud coverage, cloud height, cloud radiance, and surface albedo, are test constants in this study. There are three contributing steps that most influence the discrepancy between retrieval products: stratospheric separation, atmospheric radiative transfer, and a priori vertical distribution of NO_2 (OMI Algorithm Theoretical Basis 2002). These three processes have profound implications that ultimately affect actual air quality control.

1.1 OMI NO_2 Algorithm Theoretical Basis

Satellite measurements of NO_2 slant columns are based on Differential Optical Absorption Spectroscopy or DOAS. The slant column density is determined by the absorption of photons along the path of light of interest to the sensor relative to the

number of photons at the top of the atmosphere. The ratio of the radiance at the top of the atmosphere, I, to the extraterrestrial solar irradiance, E, is defined as the reflectance spectrum, R: equation 1. In general, I, E, and R are functions of wavelength, zenith angle and sun-satellite azimuth angle.

$$R(\lambda) = \frac{\pi \cdot I(\lambda)}{\mu_0 \cdot E(\lambda)} \qquad (1)$$

A least squares fit of the modeled spectrum to the natural log of the measured spectrum is used to relate radiance to the NO_2 slant column concentrations. This is assumed to obey the modified Lambert – Beer law: equation 2.

$$\ln[R(\lambda)] = - \sum_i \sigma_i(\lambda) N_{s,i} - P_3(\lambda) \qquad (2)$$

This formula is solved for the NO_2 slant column density, N_s. Sigma, σ, is the effective absorption cross-section including the ring effect and the absorbing effects of other molecules plotted in Figure 1.1.1 for each NO_2 molecule, i. The third-order polynomial, P_3, accounts for multiple scattering and absorption, Mie scattering, and surface albedo. The slant column densities and polynomial coefficients are determined by a least square fitting that minimizes the cost function between the observed reflectance and the modeled reflectance.

A fitting window with wavelengths of 405 – 465 nm is used by KNMI and NASA to minimize NO_2 slant column concentration errors (Figure 1.1.2) and to avoid absorption lines from other chemical species. This window is selected by these two groups because it circumvents the ring effect related to the Ca(II), H, and K absorption lines at 397 – 393 nm and O_2–O_2 absorption lines at 467 – 475 nm. Reference spectra for the absorption

cross section of NO_2 were obtained from *Vandaele et al.* [1998]. O_3 cross sections were obtained from *Bogumil et al.* [1999], H_2O cross sections from *Harder and Brault* [1997], O_2–O_2 cross sections from *Newnham and Ballard* [1998], and ring effect spectra from *Chance and Spurr* [1997].

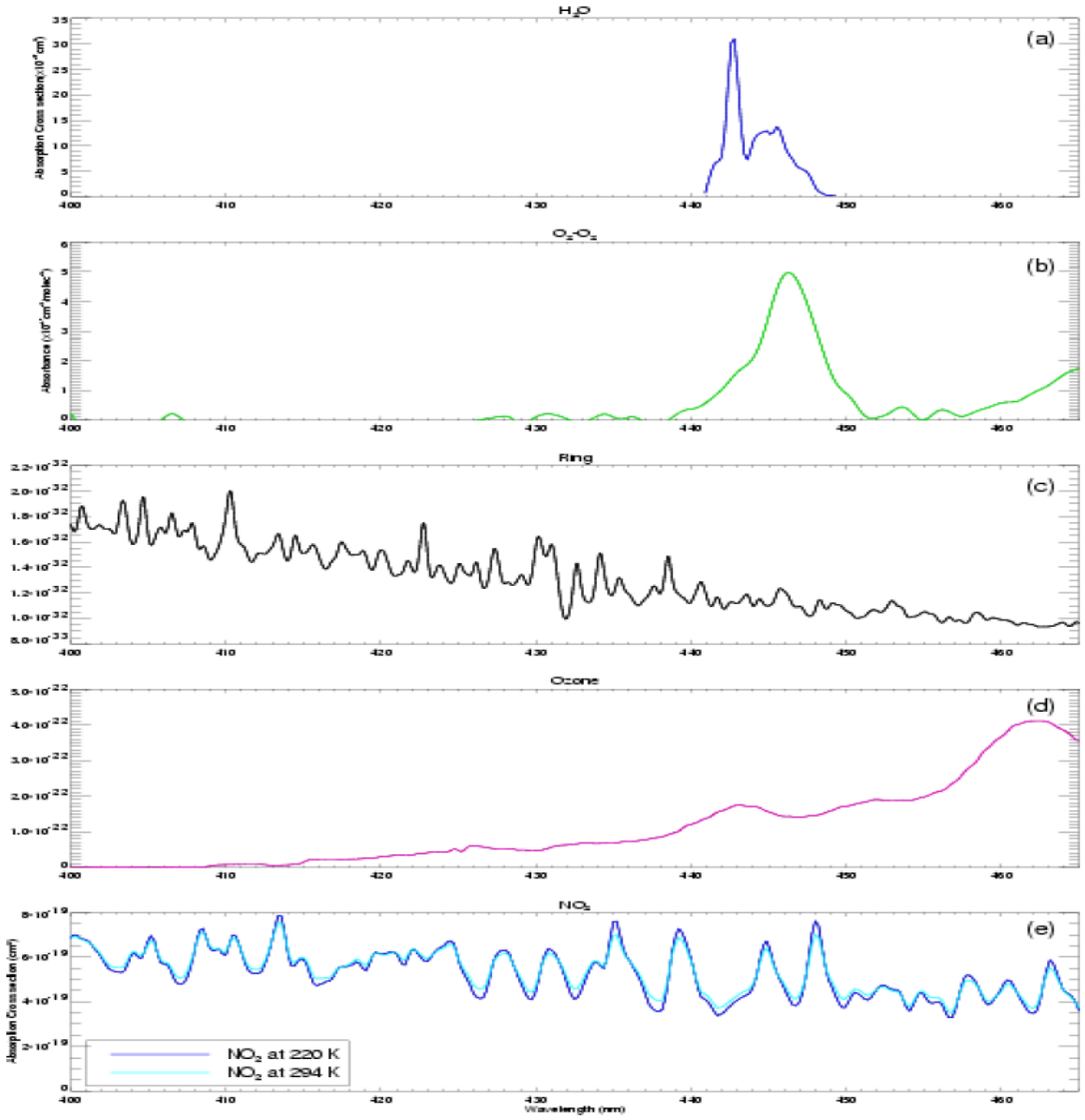


Figure 1.1.1: The major absorbers within the OMI NO_2 window. Reference spectra were used to calculate NO_2 slant columns from OMI satellite observations: (a) H_2O, (b) O_2-O_2, (c) Ring effect, (d) O_3 (e) NO_2 [OMI Theoretical Basis 2001].

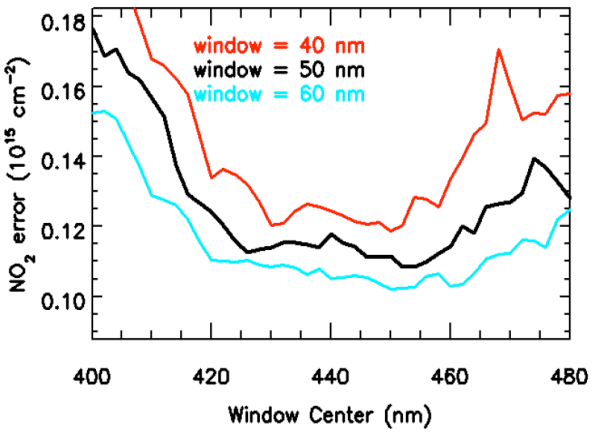


Figure 1.1.2: The fitting window currently used by TEMIS to look for NO_2 in the slant column. Each of the three lines shows the fitting error as a function of wavelength for a given window width. A width of 60 nm is currently used by TEMIS.

Air mass factors are used to estimate vertical columns from the observed slant columns. The air mass factor is defined as the slant column concentration divided by the vertical column concentration. This is computed for each level of the chemical transport model. The air mass factor, M, then is the integral of the altitude-resolved air mass factor, m, times the NO_2 temperature correction factor times the independent volume density profile, n.

$$M(z) = \frac{\int_z^\infty m(z') \cdot \alpha\left[T(z'), T_0\right] \cdot n(z') dz'}{\int_z^\infty n(z') dz'} \qquad (2)$$

The independent volume density profile in this case is NO_2 predicted by the a priori chemical transport model. The air mass factor is affected by the vertical distribution of NO_2 in the chemical transport model and not the magnitude of the model. The viewing geometry of the satellite determines the angle and length of the slant column which influences the altitude-weighted air mass factor, m, equation 2. The geometric air

mass factor is a function of solar zenith angle, viewing zenith angle, and relative sun-satellite azimuth angle. If, for example, the satellite is looking straight down and the sun is directly behind the satellite, then the geometric air mass factor would be exactly two.

Two other complications affecting the altitude weighted AMF are clouds and surface albedo. These parameters act to scatter more or less photons into the light path. The air mass factor for a partly cloudy pixel, M', is derived from the independent pixel assumption (3). Where, $M(z_c)$ is the air mass factor above cloud-top height, and $M(z_o)$ is the air mass factor above the ground for a clear scene. The independent pixel approximation assumes that air mass factors can be written as the linear combination of cloudy and clear scenes.

$$M' = w \cdot M(z_c) + (1 - w) \cdot M(z_0) \qquad (3)$$

The radiance-weighted cloud fraction, w, is a function of the OMI effective cloud fraction, c, obtained from FRESCO and the irradiance of the cloud, I_{cloud}, and the irradiance of the ground, I_{clear}.

$$w = \frac{c \cdot I_{cloud}}{c \cdot I_{cloud} + (1 - c) \cdot I_{clear}} \qquad (4)$$

At nadir, the resolution of OMI is 13 x 24 km^2, and as the viewing zenith angle increases, resolution decreases to, about 40 x 160 km^2 on the ends [*Boersma et al. 2008*]. The direction of the flight and the orientation of the sensor onboard OMI are not perpendicular (*Figure 1.1.3*). This asymmetry gives different viewing geometries and cell resolutions on either side of the satellite. Since NO_2 retrievals are a function of viewing geometry, this asymmetry is reflected in the retrieval product.

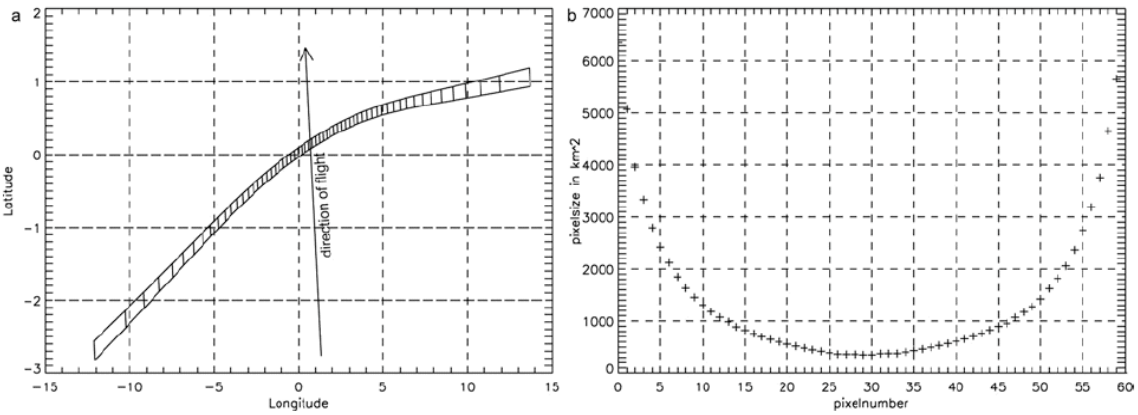


Figure 1.1.3: 60 pixel cross track of OMI, geometry and area [*Boersma et al.2008*].

1.2 Differences between NASA and KNMI Algorithms

There are currently several differences between NASA's retrieval and KNMI's retrieval. These differences arise entirely from converting OMI slant columns into vertical columns and tropospheric vertical columns. Aside from a priori profiles, there are two retrieval algorithm categories in which NASA and KNMI differ: (1) the stratosphere–troposphere NO_2 separation and (2) the degree to which the retrieval is sensitive to near surface NO_2.

After air mass factors are used to convert slant columns to vertical columns, the stratospheric NO_2 contribution must be subtracted from the vertical column to get the tropospheric column. NASA stratosphere – troposphere separation works by first assuming that there is no tropospheric NO_2 over the ocean. Such vertical columns containing only stratospheric NO_2 are called "clean columns". Wave analysis is then applied to these clean columns to derive the global stratospheric contribution. This stratospheric model is then subtracted from the "polluted" columns to get the

tropospheric contribution. KNMI uses a global chemical transport model to predict tropospheric and stratospheric contributions. KNMI assumes its stratospheric model from TM4 and subtracts that directly from the retrieved vertical columns to get the tropospheric contribution.

KNMI and NASA treat sensitivity to near surface NO_2 differently as well. This discrepancy is a function of the NO_2 radiative transfer model. Another difference between KNMI and NASA is how the a priori profiles are computed. NASA uses an annual average NO_2 a priori profile, whereas KNMI uses a fully-coupled retrieval/chemical transport model. The operation flow chart for each algorithm is discussed in Figure 1.2.1. After each swath the analysis from the retrieval updates the chemical transport model's initial conditions and is rerun with those conditions [Boersma et al., 2007].

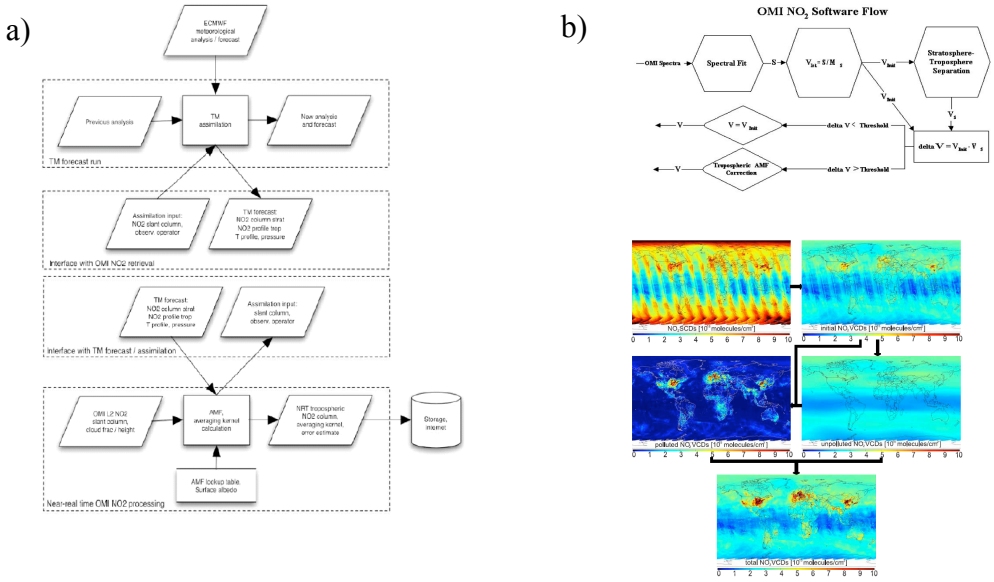


Figure 1.2.1: KNMI near-real-time and NASA standard product flow charts. KNMI chemical transport model and retrieval coupling (a) and flow chart and graphic representation of the NASA Standard Product (b)

The differences between the NASA OMI standard product and the KNMI OMI near-real-time product are summarized in Table 1.2.1. Both NASA OMI and KNMI OMI determine the observed slant column density from a spectral fit to Earth's reflectance spectrum as described above. However, for the stratospheric–tropospheric separation of the slant column, the two scripts begin to differ [Boersma et al., 2007]. KNIM uses a retrieval–modeling chemical transport model to calculate the stratospheric slant column contribution. This stratospheric contribution is subtracted from the observed slant column to give the tropospheric residual slant column. This tropospheric slant column is then scaled by the inverse of the tropospheric air mass factor, a function of the a priori and radiative models, to give the tropospheric vertical column. NASA OMI separates out the stratosphere by first assuming ocean tropospheric columns to be clean of NO_2. The NO_2 measured is assumed to be entirely stratospheric, on which a global wave-2 zonal band fit is applied. This is subsequently used to calculate the tropospheric slant column residual. Air mass factors are calculated using yearly averaged NO_2 profile shapes from GEOS-Chem and NASA's own radiative model [Boersma et al., 2007]. In this thesis KNMI TM4 and GT REAM a priori models and NASA TOMRAD and KNMI DAK radiative models are compared.

Table 1.2.1: Resolution and retrieval limits of the radiative transfer look-up-table

	Pressure	SZA	VZA	RAA	Albedo	Surface Pressure
DAK – resolution	24	13	9	10	10	10
TOMRAD – resolution	51	5	4	4	5	5
DAK – high limit	1100	90	60	180	1	1100
TOMRAD – high limit	1013	88	70	180	1	1013
DAK – low limit	0	0	0	0	0	0
TOMRAD – low limit	0.8097323	0	0	0	0	55.96162

1.3 - Error analysis of satellite algorithms

Other than these a priori NO_2 chemical transport model sensitivity differences between NASA and KNMI, there are several other corrections which were made to the underlying retrieval scheme. These include a temperature correction, a cloud fraction correction, and corrections for satellite specific problems such as systemic across-track errors. The NO_2 cross-section fit is significantly influenced by the temperature of the absorbing NO_2. Between the warmest and coolest atmospheric NO_2 temperatures, this difference may be as much as 15%. Figure 1.3.1 shows the amount of error in the slant column density that would occur if temperature was not accounted for. In the extreme North Atlantic, NO_2 slant columns would be overestimated by as much as 10%. Whereas NO_2 slant columns over the Northeastern United States and Western Europe would be underestimated by as much as 10%. The ECMWF six-hour forecast temperatures are used in the TM model. Errors in the ECMWF temperature errors are accurate up to a few degrees Kelvin, which translates to as much as a 1% error in the NO_2 slant column [Boersma, 2004].

Table 1.3.1: Summary of differences [Boersma et al., 2007]

	OMI standard product (Bucsela et al., 2006)	OMI near-real time (Boersma et al., 2007)
Slant column	DOAS (405–465 nm)	DOAS (405–465 nm)
Across-track variability correction	Correction factors based on 24-h data[1]	Correction factors computed per-orbit[2]
Stratospheric slant column	Wave-2 fitting along zonal band	Data assimilation in TM4
AMF - cloud parameters	O_2-O_2(Acarreta et al., 2004)	O_2-O_2(Acarreta et al., 2004)
AMF - surface albedo	GOME(Koelemeijer et al., 2003)	TOMS/GOME[3]
AMF - profile shape	Yearly average profile shapes from GEOS-Chem ($2.5°\times2°$)[4]	Collocated daily output at overpass time from TM4 ($3°\times2°$)
AMF - ghost column	Not included	Implicit in AMF definition

[1] In the standard product, mean slant columns are adjusted for a given cross-track position to the mean value at all positions. The mean slant column and the mean initial AMF for each cross-track position are computed from 24-h of data using measurements obtained at latitudes between ±55°. These are used to generate a set of 60 correction constants, one for each cross-track position, which are subtracted from the slant column values before computation of the vertical columns.
[2] The correction for the near-real time retrieval is described in Sect. 4.1.
[3] Surface albedo fields are taken from a combination of Herman and Celarier (1997) and Koelemeijer et al. (2003) as described in Boersma et al. (2004).
[4] AMFs are corrected based on average a priori profile shapes when the retrieved slant column is larger than the estimated stratospheric slant column (Bucsela et al., 2006).

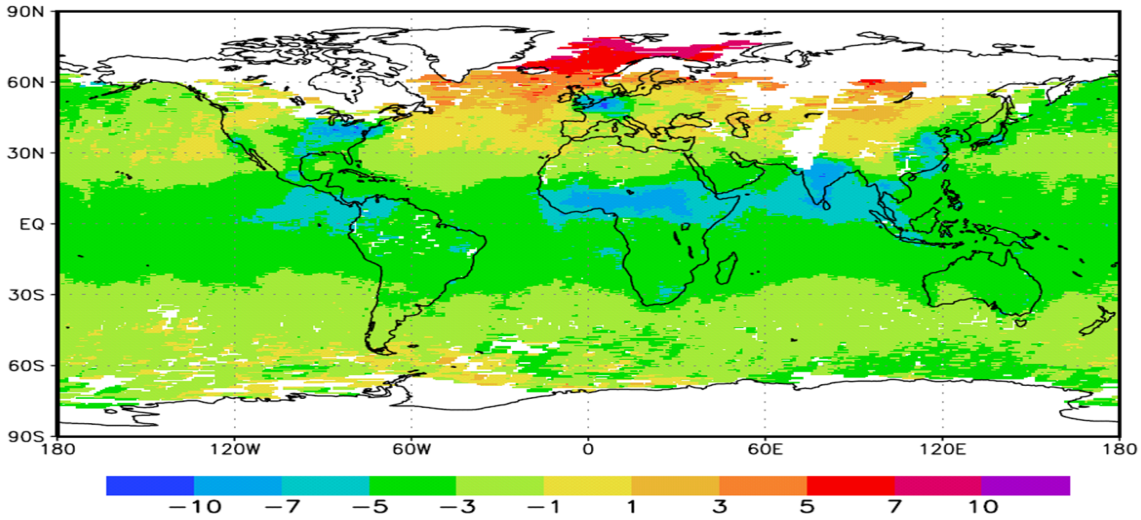


Figure 1.3.1: The percent effect of a priori temperature [Boersma et al., 2007].

Lambert-equivalent ground reflectivity is assumed in the retrieval and has a large impact on AMF calculations. Albedo uncertainty is about 2% from 60S to 60N and is larger outside this range. For example, if a slant column is partially influenced by snow, yet climatology predicts no snow, systematic errors may occur. The KNMI cloud radiance correction is thought to correct for this snow bias by creating an effective cloud at the ground that increases the cloud radiance fraction [*Boersma et al., 2004*].

As a result of these known differences, resulting in the known total AMF Uncertainty (Figure 1.3.2), one would expect that NASA and KNMI retrievals would react differently to the high resolution a priori chemical transport model REAM. As described in the methods, to avoid confusing results and to allow for comparison of a priori models and radiative models, only certain retrieval methods are used.

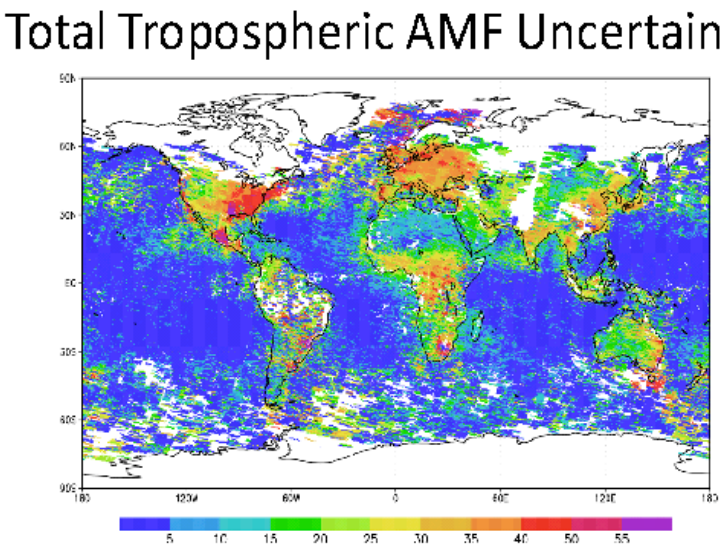


Figure 1.3.2: Total tropospheric air mass factor uncertainty accounting for cloud fraction, cloud height, surface albedo, and profiles shape with cloud radiance fractions of less than 0.5. [Boersma K. F. 2004].

Table 1.3.1: Results Sensitivity Analysis of the OMI DOAS NO2 algorithm [*OMI, 2002*].

Slant column errors	**Unpolluted case error (%)**	**Polluted case error (%)**
NO_2 cross section	2	2
Temperature	4	6
Instrument noise	4	2
Spectral calibration	0.5	0.3
Air mass factor errors		
NO_2 profile shape	1	20
Surface albedo	0.5	20
Cloud albedo	0.5	4
Cloud fraction	0.3	8
Cloud pressure	0.5	50
Aerosol assumption	1	15
Error in $N_{v,u}$		
Estimation of $N_{v,u}$	N/A	5

Table 1.3.2: Estimated accuracy of the OMI DOAS NO2 algorithm [*OMI, 2002*].

Vertical column errors	**Unpolluted case Total column error**	**Polluted case Total column error**	**Polluted case Tropospheric column error**
Clear	5 %	20 %	30 %
Partly Cloudy	5 %	50 %	60 %

1.4 Validation of KNMI and NASA

Satellite retrievals are not direct observations, but rather an observation weighted by an a priori model. This model is used to determine vertical distribution of NO_2 and is scaled by an atmospheric radiative model to back out the corrected magnitude of the observation. That magnitude is constrained by performing a radiative balance at the top and throughout the atmosphere. The radiative and a priori models are then verified against an independent source of measurements for robustness and accuracy. Several organizations have completed tests to check satellite retrievals to assure they are functioning properly [Ionov et al 2008].

Wenig et al. [2004] used brewer data to validate NASA standard product retrieval. This data was compared to OMI cells within 15 km (the red line) and 50 km of the site (the blue line). Figure 1.4.1 shows that brewer data collected to measure NO_2 vertical column concentrations actively demonstrates that the NASA standard product consistently underpredicts NO_2 vertical columns by as much as 50%. Celarier et al. [2008] found very similar results using the same technique. Celarier et al. [2008] report that NO_2 can vary significantly over tens to hundreds of meters and that, to some degree, ground measurements should be higher than satellite retrievals due to the averaging effect of the area observed by the satellite.

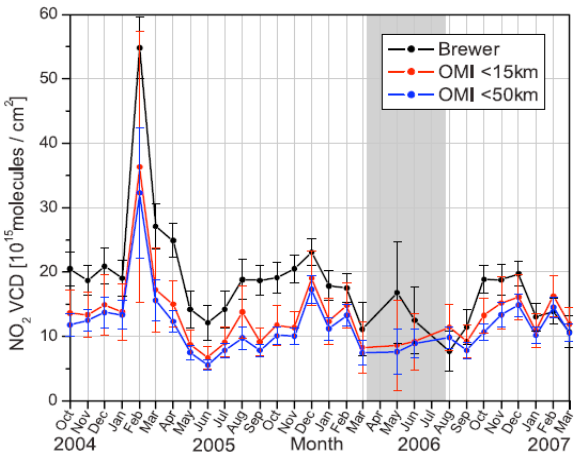


Figure 1.4.1: Black, brewer data collected near OMI pass time, NASA Goddard Space Flight Center (38.98N, 76.83W, 90 m above sea level) [*Wenig et al., 2004*].

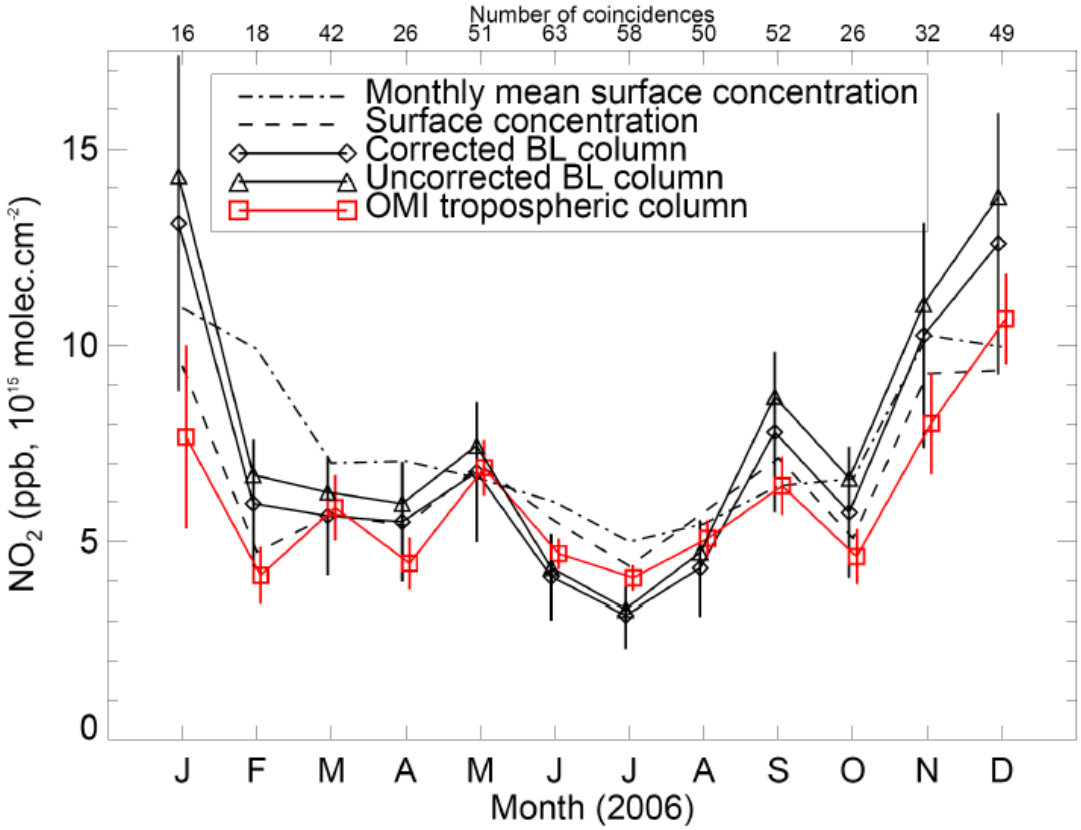


Figure 1.4.2: Comparison of KNMI OMI Near Real Time Product to SCIAMACHY over Israel [Boersma et al., 2009]

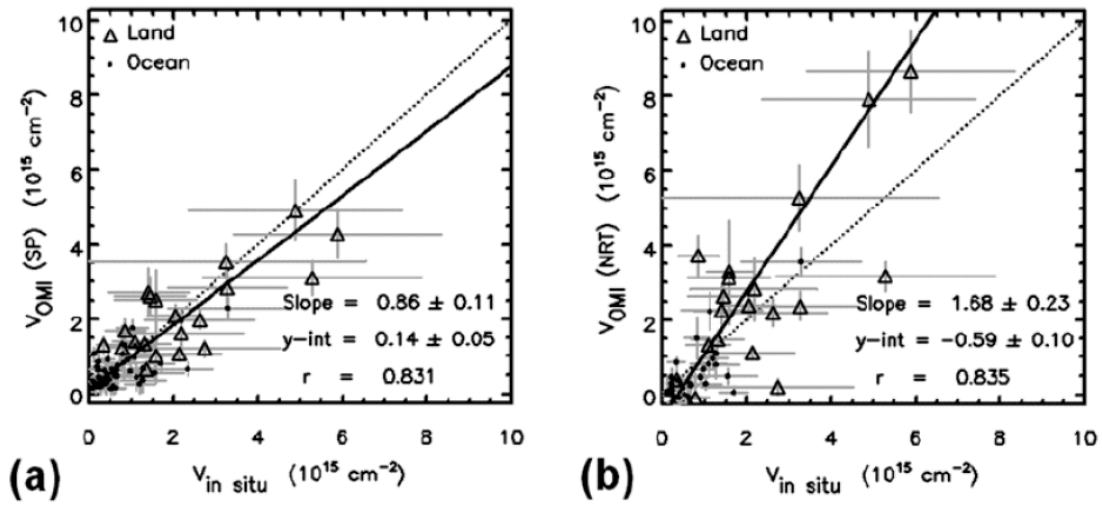


Figure 1.4.3: Direct comparison of NASA and KNMI using aircraft data from the INTEX-B mission [Bucsela et al. 2008].

KNMI OMI Near Real Time Product is compared with SCIAMACHY and boundary layer estimates over Israel in Figure 1.4.2. Here, KNMI OMI agrees reasonably well with ground measurements [Boersma et al., 2009]. The paired t-test did not find that either KNMI near-real-time or the NASA standard product exhibited significant differences between the in-situ and corresponding satellite measurements [Bucsela et al. 2007]. However, the INTEX-B mission demonstrates plainly that KNMI OMI overestimates NO_2 vertical columns to a larger degree than NASA underestimates it.

1.5 NO_2 Retrieval Trends, Inversions and Variability

To evaluate the trend of NO_2 emissions, Van der A et al. [2008] created a time-series of global NO_2 emissions from a session of several inversions. Only cells from 60S to 60N were used and those cells with high albedo or high solar zenith angles were

ignored. Pixels with a statistically significant trend apart from zero are plotted in Figure 1.5.1. East China and parts of West Asia show a strong significant trend in increasing NO_2 emissions. Other locations such as in the Ohio Valley and northwestern Europe have reduced emissions as a result of strict regulation on power plant emissions. Nevertheless, much of North America and Europe remain statistically indifferent. For other regions where less significant trends are observed, changes in albedo and micro-climate could also affect NO_2 inversions, creating trends which are not necessarily the true story. Martin et al, showed similar emissions inventories using GOME [Martin, R. V., 2003].

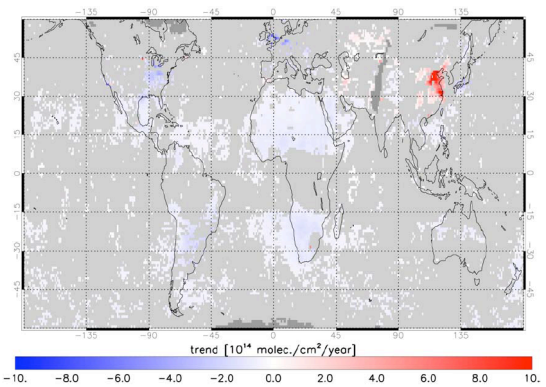


Figure 1.5.1: Linear trend in NO_2 emissions.

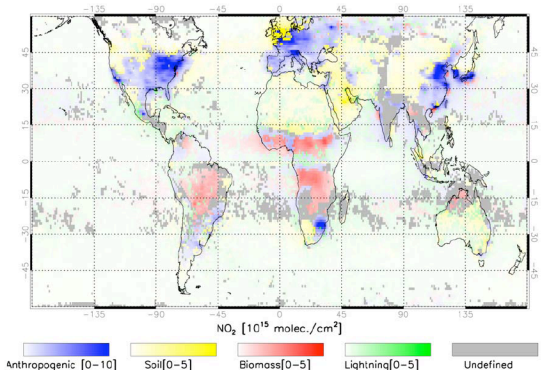


Figure 1.5.2 – The dominant NOx source

The distinction between the sources of NO_2 (Figure 1.5.2) is arrived at by comparing the mean NO_2 columns over cloudy pixels and completely clear pixels at the same location. If NO_2 is correlated with the cloudy pixels, then lightening is said to be the dominant source. The difference between soil emissions and biomass burning is determined by the month of peak NO_2. Soils typically have their maximum in the summer, and if not, it must be biomass burning. Anthropogenic NO_2 is determined by finding the variability of the NO_2 retrieval column. Biomass burning is highly variable, whereas most anthropogenic sources are constant. By making some assumptions about the variability of daylight and ignoring meteorological issues, the anthropogenic dominant source can be back calculated [Van der A et al., 2008].

The REAM model, as described in detail by C. Zhao et al [2009], is a tropospheric regional chemical transport model consisting of 23 vertical layers under 10 hPa and a horizontal resolution of 70 km. Figure 1.5.3 shows REAM compared to SCIAMACHY retrievals for four different simulations. One was run with MM5, one was run with WRF, and each was run with a sensitivity test to lightning. Lightning is an important contributor of NO_2 in the troposphere and is important in determining the distribution of NO_2 in the a priori NO_2 profile of the retrieval. WRF–REAM with lightning was chosen since it best agreed with the INTEX-NA mission, (Figure 1.5.4). Other model simulations significantly differed from the observations.

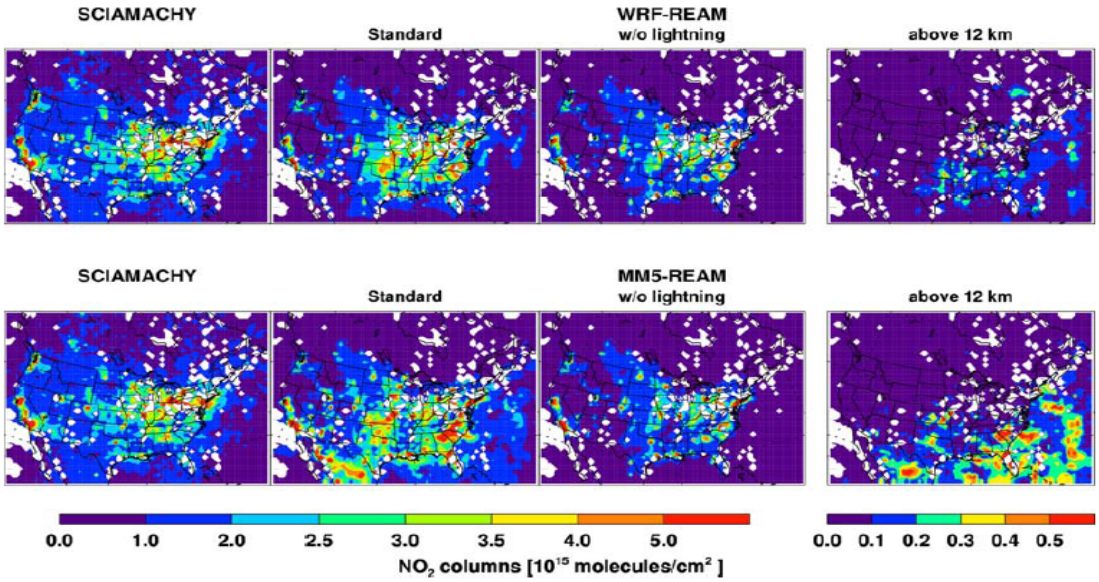


Figure 1.5.3: NO_2 SCIAMACHY retrievals compared to REAM simulations using four different simulations [Zhao et al., 2009].

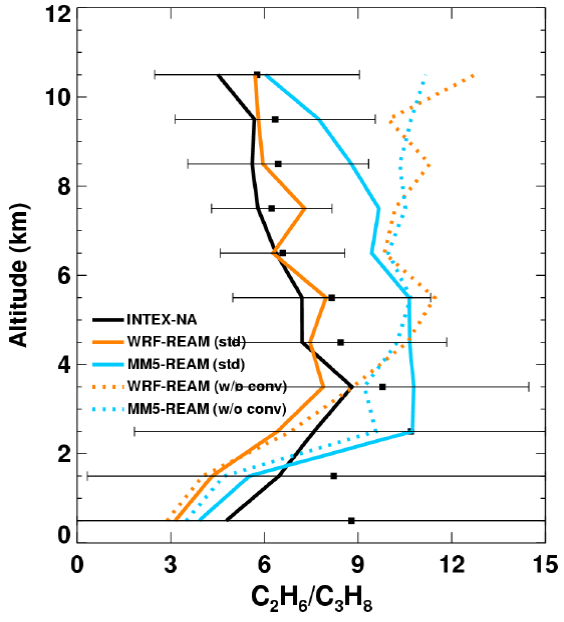


Figure 1.5.4: REAM simulations compared to INTEX-NA mission

1.6 Purpose for the proposed research

A review of the literature presents many scientific questions remaining between the difference of the retrieval data products KNMI and NASA. The purpose of this project is to verify the claims made previously in the literature as presented in this introduction by identifying new sources of the discrepancies. Some of these questions include: why does the NASA product produce a smaller vertical column as opposed to KNMI and how does an a priori chemical transport model affect the retrieval. Is there in fact no statistical difference between the two retrievals using these different models' outputs? The comparison uses f-tests and t-tests to assess distribution and means respectfully of NO_2 tropospheric vertical column output with a 95% confidence interval.

This approach justifies the comparison of the most significant contributors to the retrieval process. Other, less significant contributors are all held constant for this evaluation. These include parameters like cloud pressure, cloud radiance, cloud albedo, cloud faction and ground albedo. In brief, KNMI is found to be more sensitive to near surface NO_2 because its radiative transfer model weighs NO_2 at this location disproportionately more than NASA. The a priori models play a significant role in the retrieval. Since these a priori models determine the NO_2 distribution in the slant column, they dictate, in effect, what the satellite observes. Depending on where that NO_2 is located in the slant column, the radiation model determines a proportionately greater or smaller vertical column.

CHAPTER 2

METHODS

A variety of methods were applied to test for consistency and robustness of both the KNMI and NASA radiation models and both of the TM4 and REAM a priori models. Currently, tests are broken down to analyze the two of the three major contending parts of the NO_2 satellite retrievals. These include the NO_2 radiance in Earth's atmosphere calculation, the a priori NO_2 vertical profiles, but not the stratospheric–tropospheric separation. Figure 2.1 shows the options selectable for a particular "meshed" retrieval product.

2.1 Stratosphere–Troposphere Separation

To convert the observed slant columns at each satellite overpass to vertical columns the contribution of NO_2 in the stratosphere and troposphere must be known. This is because the satellite has less sensitive, lower observation efficiency to NO_2 in the troposphere then the stratosphere. As described in the introduction, NASA attempts to observe the daily stratospheric contribution by fitting a wave-2 fit to clean pixels where as the KNMI assumes that the TM4 chemical transport model accurately describes the stratosphere, subtracting the slant column equivalent directly from observations.

These stratospheric separations were compared in six ways. The TM4 stratosphere is used with DAK radiance and REAM a priori tropospheric profiles, DAK radiance and TM4 a priori troposphere profiles, DAK radiance and GEOS-Chem profiles. Likewise

NASA's unpolluted stratospheric separation is used with same above three options. The relative difference is computed and discussed in the results section.

2.2 NO_2 Constant Profile Radiance Test

KNMI and NASA run two different off-line radiance models to calculate the sensitivity of the satellite to NO_2 molecules under different environmental conditions. These models produce a lookup table that is a function of satellite geometry (the solar zenith angle, the viewing angle) and the relative sun-satellite azimuth angle and a function of the environment (pressure and albedo).

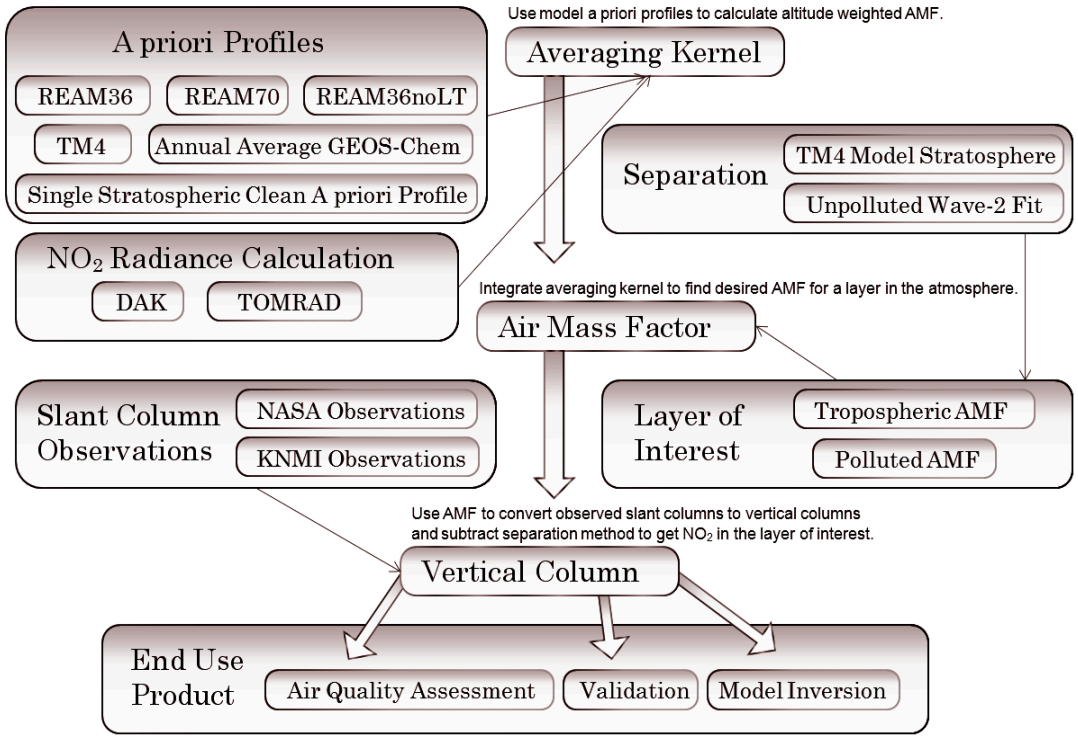


Figure 2.1: NO_2 Retrieval Flow Chart

The constant profile test employs one single a priori profile over the continental United States region. The aim is to quantify the difference in vertical column calculations between NASA and KNMI scripts as result of their respective NO_2 radiance model, TOMRAD and DAK, and not the a priori profiles. A series of eight specific test profiles were developed to quantify satellite sensitivity to the radiance model. Four profiles were selected exactly from TM4 OMI data: a clean profile (no tropospheric NO_2) and three increasingly dirty profiles. These profiles were selected to observe the sensitivity of NO_2 near the surface compared to NO_2 distributed throughout the tropospheric column. They were chosen with the condition that they have similar stratospheric vertical columns. In addition, four profiles were created artificially. Each of these had the same vertically integrated tropospheric vertical column: 3.6E15 molecules per square centimeter and exactly the same stratospheric distribution and tests for NO_2 concentrated at different pressure altitudes. All eight profiles were run over North America for month the July 2006.

2.3 A priori vertical column comparison

Three a priori models (TM4 provided by KNMI, GEOS-Chem provided by NASA, and the REAM provided by the department of Earth and Atmospheric Science at Georgia Tech) are used as the a priori input to determine NO_2 vertical distributions in the troposphere. GEOS-Chem annual average profiles are not presented in this thesis. Only REAM and TM4 hourly profiles were used. These two models consisted of several structural differences. REAM70 is a regional tropospheric model with 23 layers, a spatial resolution of 70 km, and uses WRF meteorology. REAM36 has a spatial resolution of

36 km. TM4 is a global tropospheric, stratospheric model with 34 layers, a spatial resolution of 1 degree and uses ECMWF meteorology.

The a priori profiles were necessary to initialize the retrieval with a vertical distribution of NO_2. This distribution is applied to the radiative transfer model to determine NO_2 vertical column sensitivity for each layer of the a priori vertical profile. This sensitivity factor scales the satellite observation up or down accordingly and is termed the altitude-weighted-AMF or averaging kernel. This averaging kernel is integrated vertically to determine retrieval layer AMFs and then vertical columns.

This process is limited by the vertical resolution of both the a priori vertical profiles and the vertical resolution of the radiative transfer model. KNMI and NASA radiative transfer models have the following resolutions and ranges for the offline parameters: REAM70 km has 23 vertical layers in the troposphere and REAM36 has 36 vertical layers in the troposphere. TM4 has about 20 layers in the troposphere and 14 layers in the stratosphere. This gives REAM70 37 layers total, REAM36 50 layers total, and TM4 34 layers total. REAM36 and NASA TOMRAD are about equivalent in vertical pressure resolution. In the retrieval, the radiative transfer model is linearly interpolated to fit the a priori vertical profile.

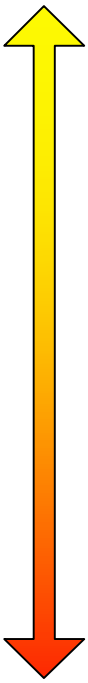
KNMI interpolates TM4 model output for each OMI satellite cell. To work with the horizontal resolution differences between REAM and TM4, with each satellite cell, the nearest REAM index is used. Each satellite cell is gridded back to REAM resolution

and weighted by the cosign of the viewing zenith angle. This results in different TM4 averaging grids, TM470 and TM436. Several steps are taken to solve the vertical differences between TM4 and REAM. Any levels above TM4 tropopause on REAM are added to the REAM troposphere and the TM4 full stratosphere is copied over to REAM. This results in a pressure overlap between stratosphere and troposphere but this gives TM4 and REAM a priori models identical stratospheres. For comparison purpose only, levels are linearly interpolated from 1050hPa to 0hPa with a step size of 5.5hPa. Levels below the surface get a value of NaN. Air mass factor computations are not subjected to this vertical interpolation. The following parameters were always held constant in all case studies: cells with cloud a fraction greater than 30% are ignored, cloud albedo of 80% is used, OMI satellite geometry.

2.4 Statistical Testing on Cities in July

After all the direct comparisons are made between the different retrievals, the specific impacts on select cities of interest are analyzed to measure how well the model and the retrieval agree. Cities of interest in North America include: Atlanta, Chicago, Los Angeles, Houston, and New York City. Statistics on each city are computed over a city domain. This domain included all REAM grid points on a roughly 1 degree or smaller box centered at the middle of the city. For REAM36 retrieval a much smaller domain is used. A few of the retrievals distributions below are considered for testing. In addition, each of these retrievals are split into odd and even day distributions as described next.

- Distributions
  - **KNMI-DAK-TM4** **Actual KNMI Retrieval Product**
  - KNMI-DAK-REAM
  - KNMI-DAK-GEOSa
  - KNMI-TOMRAD-TM4
  - KNMI-TOMRAD-REAM
  - KNMI-TOMRAD-GEOSa
  - NASA-DAK-TM4
  - NASA-DAK-REAM36
  - NASA-DAK-GEOSa
  - NASA-TOMRAD-TM4
  - NASA-TOMRAD-REAM36
  - **NASA-TOMRAD-GEOSa** **Actual NASA Retrieval Product**



  - MODEL-REAM (36km, 70km, 36km nL)
  - MODEL-TM4
  - MODEL-GEOSa (‘a’ stands for annual average)

[Key: stratospheric separation – NO_2 Radiative Transfer Model – NO_2 A priori Model]
The highlighted distributions were explored in this work the remaining are yet to be done.

Cities were sampled over a period of a week or more, the distribution of tropospheric NO_2 vertical columns on even days and odd days should appear to be the same. If they are not, the retrieval or model fails the test. Statistics were computed for

each of the NO_2 distributions over each city. F-tests and T-tests were used to compare the similarity of the NO_2 daily distribution shape and mean, respectfully, to each of the other distributions within the domain of the city. Different retrievals should not produce significantly different results. Two retrievals failed the test if they were significantly different. A 95% confidence is used to test for an agreeable similarity.

The tests were performed as follows:

First, odd and even days were computed.

- KNMI-DAK-TM4-odd vs. KNMI-DAK-TM4-even
- KNMI-DAK-REAM36-odd vs. KNMI-DAK-REAM36-even
- KNMI-TOMRAD-TM4-odd vs. KNMI-TOMRAD-TM4-even
- KNMI-TOMRAD-REAM36-odd vs. KNMI-TOMRAD-REAM36-even
- MODEL-REAM36 - odd days vs. MODEL-REAM36-even
- MODEL-TM4 - odd days vs. MODEL-TM4-even

Second, retrievals were compared to one another.

- Compared the different a priori vertical NO_2 profiles
  - KNMI-DAK-REAM36 vs. KNMI-DAK-TM4
  - KNMI-TOMRAD-REAM36 vs. KNMI-TOMRAD-TM4
- Compared different radiance models
  - KNMI-DAK-TM4 vs. KNMI-TOMRAD-TM4
  - KNMI-DAK-REAM vs. KNMI-TOMRAD-REAM
- Compared the models
  - MODEL-REAM36 vs. MODEL-TM4

Lastly, compared the models with the retrievals

- KNMI-DAK- TM4 vs. MODEL-TM4
- KNMI-DAK-REAM36 vs. MODEL-TM4
- KNMI-TOMRAD- TM4 vs. MODEL-TM4
- KNMI-TOMRAD-REAM36 vs. MODEL-TM4
- KNMI-DAK-TM4 vs. MODEL-REAM36
- KNMI-DAK-REAM36 vs. MODEL-REAM36
- KNMI-TOMRAD-TM4 vs. MODEL-REAM36
- KNMI-TOMRAD-REAM36 vs. MODEL-REAM36

The results of the F-test were calculated by subtracting the test from the F-table inversion for 95% and degrees freedom of the F-test. Thus, values less than zero satisfied the null hypothesis and validated the difference in the shape of the two distributions were insignificant. In MATLAB, the formula looks similar to this: Freal = (std1/std2)^2 – finv(0.95,df1,df2), where std1 and std2 are the standard deviations of the distributions in the test and df1 and df2 are the degrees of freedom of the sample size. For the T-test, values greater than 0.05 satisfied the null hypothesis that the difference in the means of the two distributions were insignificant. In MATLAB, the formula looks similar to this: [h, p] = ttest2(data1,data2,95,'both','unequal'), where h = 1 if the null hypothesis is true, p is the statistical significance of the test, and data1and data2 are the distributions in question. In addition to f-test and the t-test, an auto covariance test is employed to check for patterns as a function of time in days within each of the twelve distributions. This indicated how the retrieval of one day is correlated with the retrieval of the next day.

CHAPTER 3

RESULTS

3.1 DAK and TOMRAD Radiance Comparison

A NO_2 radiative transfer model is necessary to determine the NO_2 effective cross-section in the column of atmosphere the satellite observes. Since OMI is a passive observer, this effective cross-section is a function of pressure, albedo, viewing zenith angle, solar zenith angle, and the relative azimuth angle as discussed in the introduction. Given these five parameters, Figure 3.1.1 shows the relative difference between KNMI and NASA's radiative transfer models DAK and TOMRAD respectively (Figure 3.1.1). This figure is the bases for all discrepancies that follow.

Figure 3.1.1 is divided into two columns. The right column is the relative difference and the left column is the absolute difference. The figure is again divided by three rows according to selective relative azimuth angles at 0, 90, and 180 degrees. An angle less than 90 degrees indicates that the retrieval is port of the satellite and conversely, an angle greater than 90 degrees indicates a retrieval starboard of the satellite. Each of the six blocks is then a function of four solar zenith angles: 0, 20, 40, and 60 on the x-axis and four viewing zenith angles: 0, 20, 40, and 60 on the y-axis. Small solar zenith angles indicate that the sun is directly behind the satellite and small viewing zenith angles indicate nadir. For each of these geometry configurations there are contour plots which are functions of albedo of zero to one and pressure from 0hPa 1000hPa on linear scales.

$$M(z) = \frac{\int_{z}^{\infty} m(z') \cdot \alpha\left[T(z'), T_0\right] \cdot n(z') dz'}{\int_{z}^{\infty} n(z') dz'} \quad , \quad (2)$$

The contours represent the difference between DAK and TOMRAD for a value of altitude resolved air mass factor, m(z'). Larger m(z') values give a larger integrated air mass factor, M(z), equation 3.1. The larger the airmass factor (M(z)) the smaller the product vertical NO_2 column. The darker colors indicate that TOMRAD computes smaller air mass factors and higher NO_2 concentrations for the same set of parameters as compared with DAK.

If a priori modeled profiles have NO_2 concentrated at high pressure, low albedo, and viewing zenith angles and solar zenith angles greater than 20 degrees, DAK will yield a smaller AMF port of the satellite but a larger AMF starboard of the satellite. This difference between port and starboard is not zero, thus leading to increased KNMI sensitivity to polluted a priori profiles near the surface on the port side of the satellite. This effect is a maximum near solar zenith angles and viewing zenith angles of 40 degrees, which are spring and fall, slightly off nadir. These patterns are demonstrated in the section 3.2 using constant model profiles over North America for the month of July.

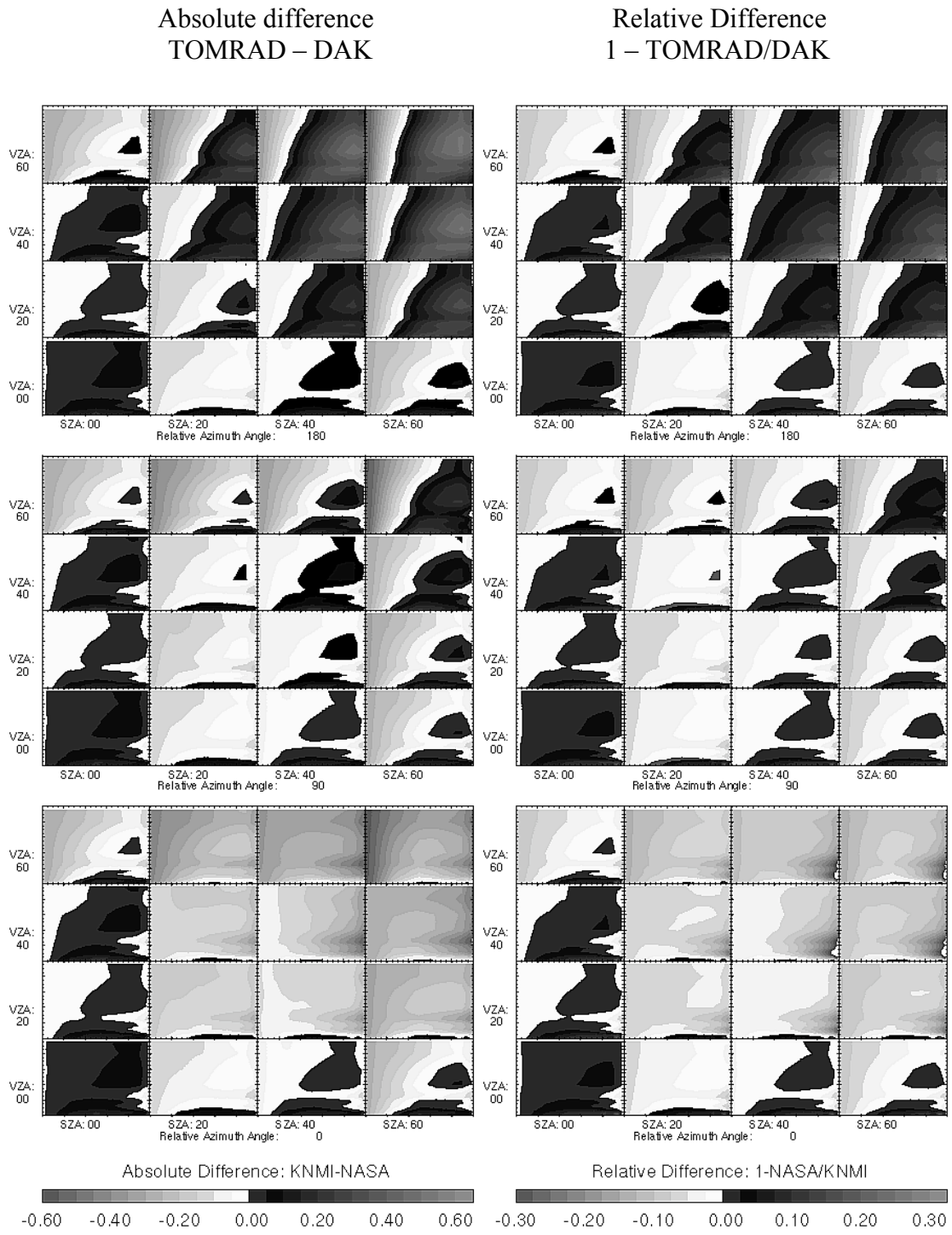


Figure 3.1.1: Differences between TOMRAD and DAK radiative models. The right column is absolute difference and the left column is the relative difference. The top row, middle, and bottom rows have a relative azimuth angle of 180 degrees, 90 degrees, and 0 degrees respectfully. The major x-axis is solar zenith angle, the major y-axis is viewing zenith angle, the minor x-axis is pressure increasing (0 to 1000hPa) and the minor y-axis is albedo from (0 to 1).

3.2 Constant Profile Results

As is discussed in the methods section, eight a priori model profiles were used to evaluate DAK and TOMRAD NO_2 radiative transfer models. The simulation over North America for July 2006 using constant values for meteorology: temperature, surface pressure, cloud fraction, and albedo and profiles, the NO_2 vertical distribution profiles, show exactly how sensitivity to NO_2 varies with retrieval geometry: relative viewing angle, solar zenith angle, and viewing angle (Figures 3.2.1, 3.2.2, 3.2.12, and 3.2.13). Environmental conditions: cloud cover, albedo, and pressure, were kept constant. The images that follow demonstrate how results found in Figure 3.1.1 relate to the realistic retrieval situation.

Figures 3.2.1 and 3.2.2 show four TM4 profiles selected to demonstrate the effect of different levels of NO_2 in the boundary layer in comparing DAK and TOMRAD radiation models. Profile 1 is the clean column case and Profile 4 is the polluted column case. Profiles 2 and 3 are intermediate. The locations from which these four profiles are selected are shown in figure 3.2.3.

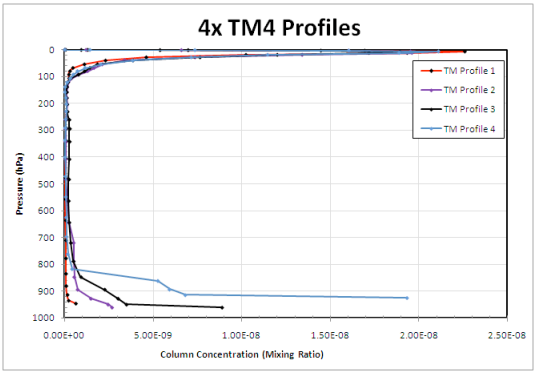


Figure 3.2.1: TM4 profile mixing ratios

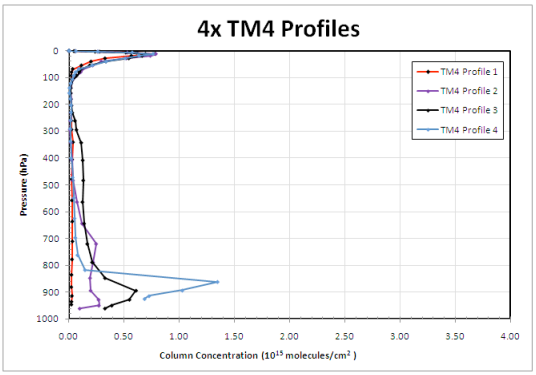


Figure 3.2.2: TM4 profile column layers

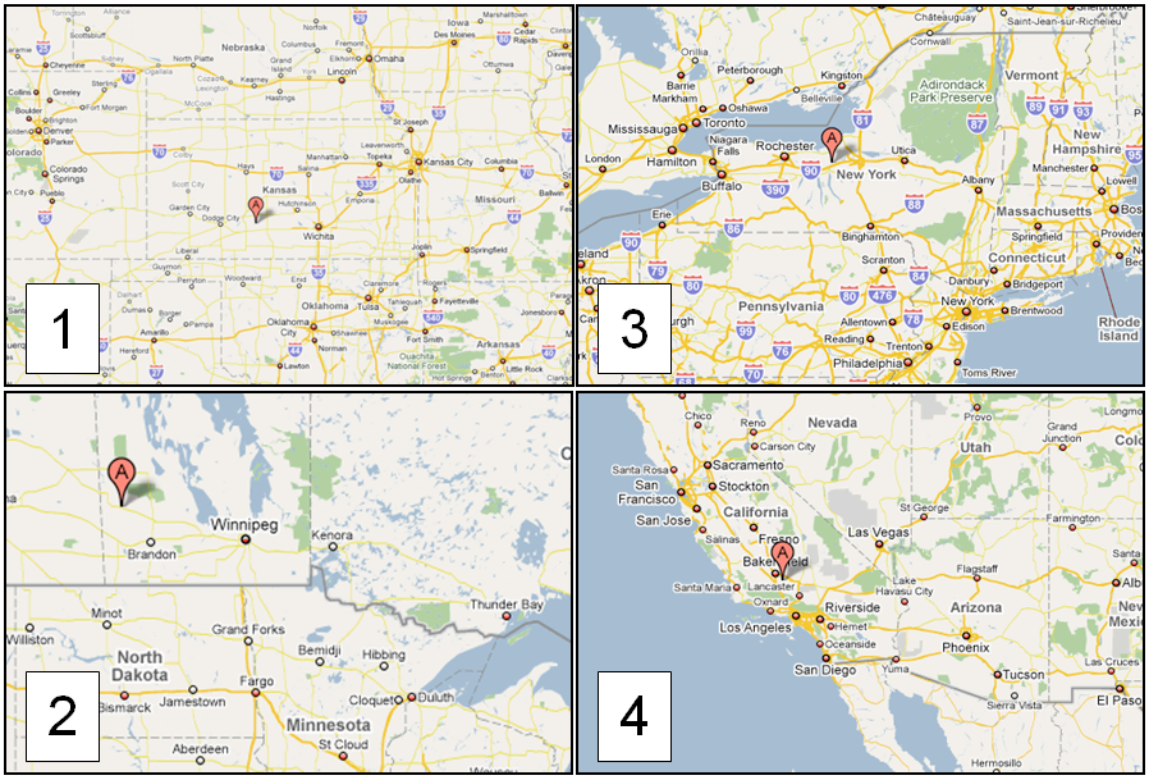


Figure 3.2.3: Location of the four TM4 profiles from Figures 3.2.1 and 3.2.2

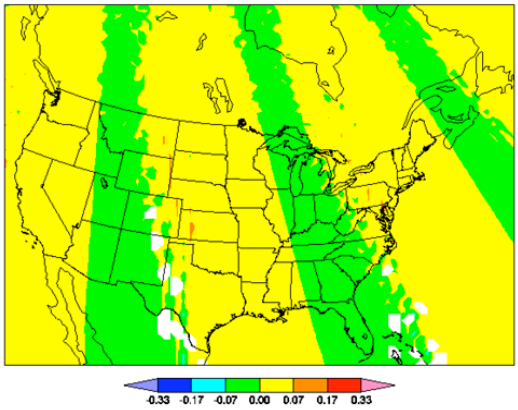


Figure 3.2.4: 1-Profile1D/Profile1T

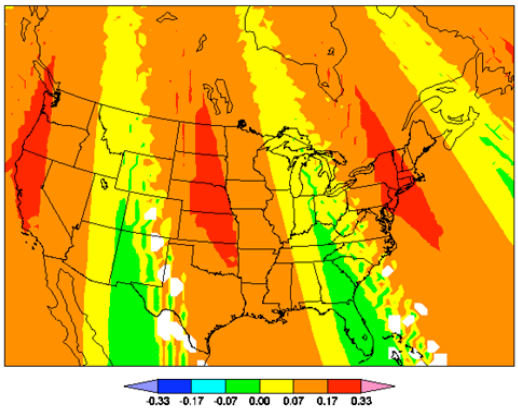


Figure 3.2.5: 1-Profile2D/Profile2T

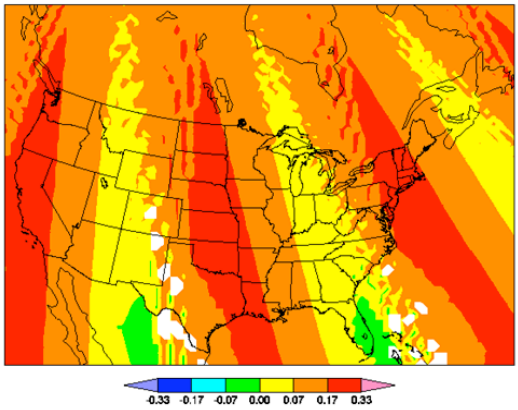


Figure 3.2.6: 1-Profile3D/Profile3T

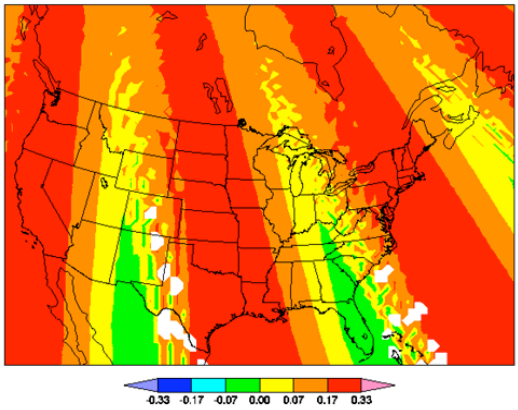


Figure 3.2.7: 1-Profile4D/Profile4T

Figure 3.2.4 through Figure 3.2.7 show the results of using a constant profile 1 through 4 over the whole domain using REAM70 grid. The white spots indicate grid points which did not receive coverage on July 1st. These grid points occur at high viewing zenith angles. Red colors indicate that DAK produces a lower integrated tropospheric AMF than TOMRAD. As pollution levels in the boundary layer increase, the difference between DAK and TOMRAD get more dramatic. In all four profiles, the relative difference between DAK and TOMRAND is positive port side and negative on the starboard side (Figure 1.1.3). The magnitude of the difference on starboard side is less than the magnitude of the difference on the port side. Hence in the duration of this paper, I refer to this issue as the DAK port side bias.

Averaging swaths from the whole month together result in Figures 3.2.8 through 3.2.11. The color bars are scaled to show spatial patterns in the test. Of the only the four profiles tested, the worst case scenario resulted in a maximum amount of port side bias of roughly 17% where everything is held constant except for the viewing geometry. This number is within the error range published by NASA and KNMI.

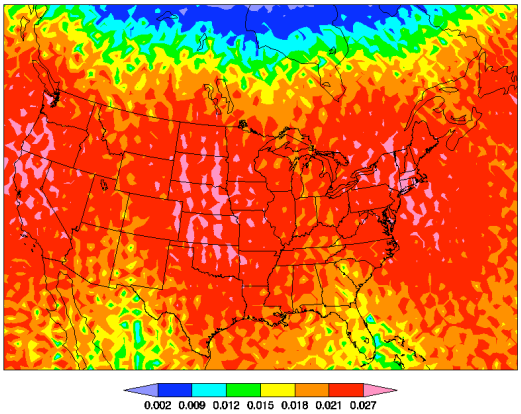


Figure 3.2.8: 1-Profile1D/Profile1T

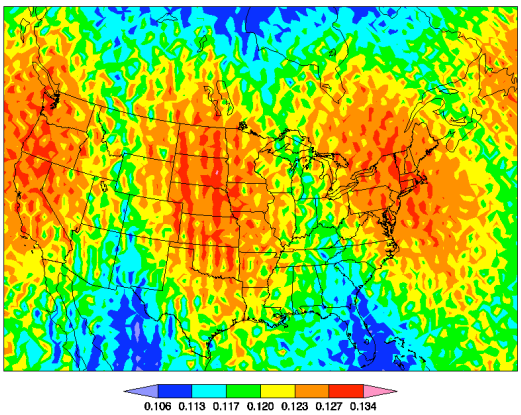


Figure 3.2.9: 1-Profile2D/Profile2T

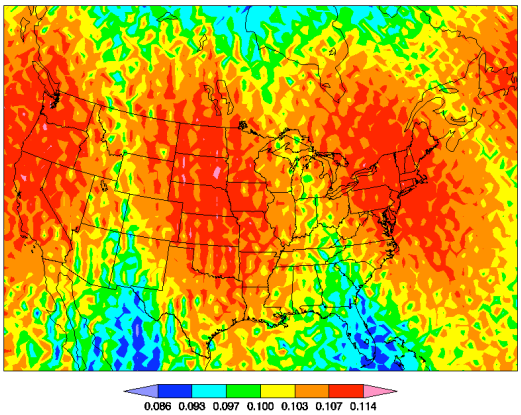


Figure 3.2.10: 1-Profile3D/Profile3T

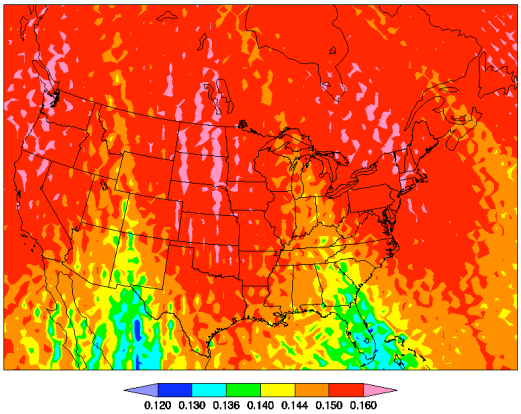


Figure 3.2.11: 1-Profile4D/Profile4T

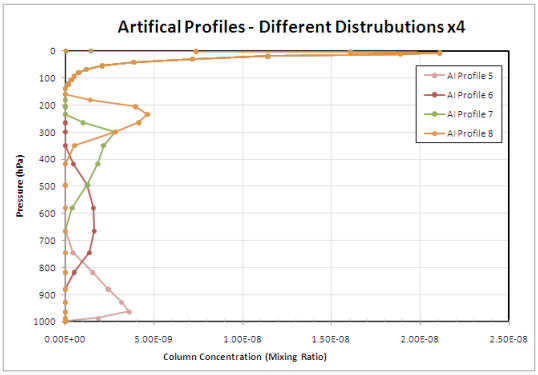


Figure 3.2.12: Artificial Profile Mixing Ratios

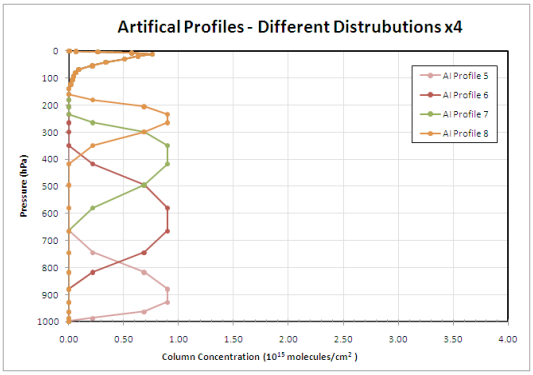


Figure 3.2.13: Artificial Profile Columns

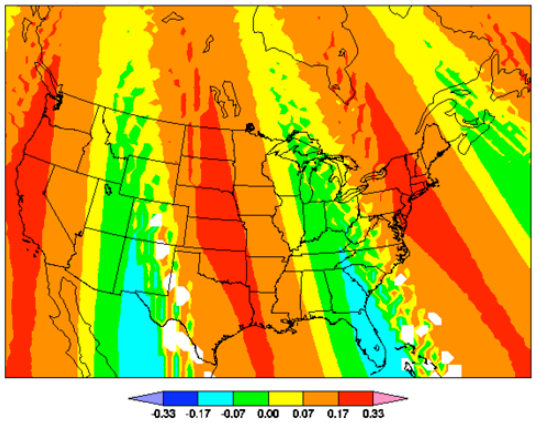


Figure 3.2.14: 1-Profile5D/Profile5T

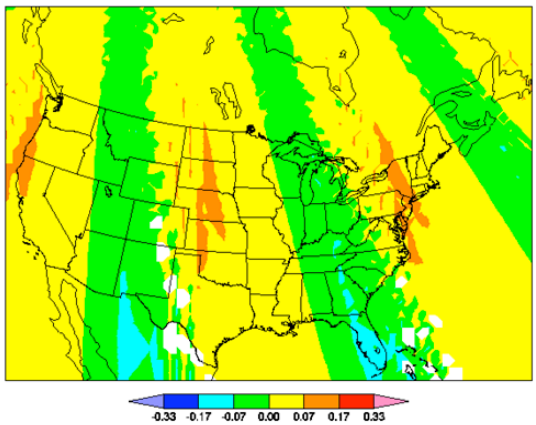


Figure 3.2.15: 1-Profile6D/Profile6T

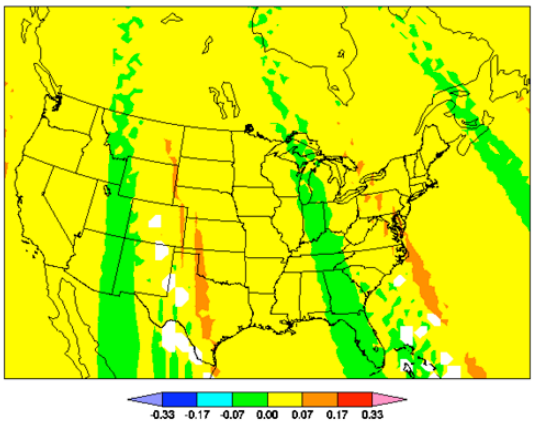


Figure 3.2.16: 1-Profile7D/Profile7T

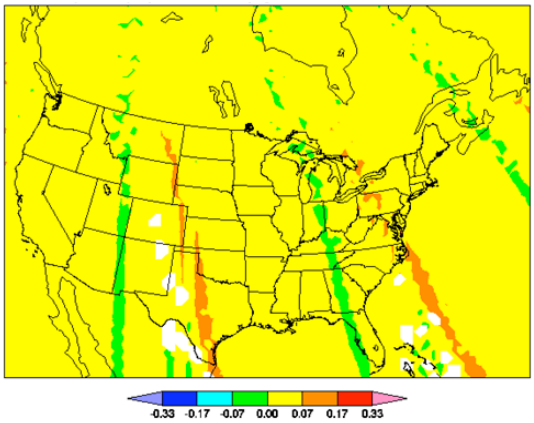


Figure 3.2.17: 1-Profile8D/Profile8T

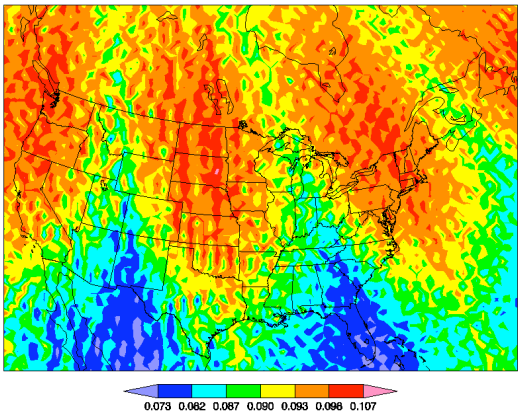


Figure 3.2.18: 1-Profile5D/Profile5T

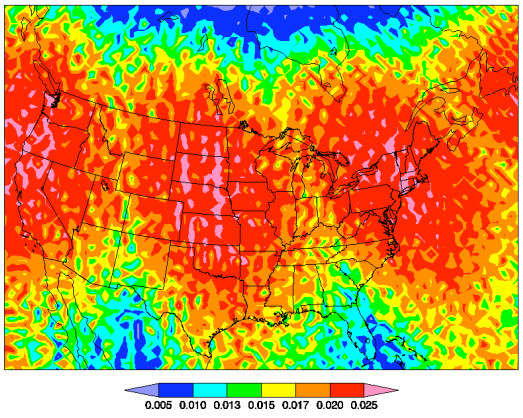


Figure 3.2.19: 1-Profile6D/Profile6T

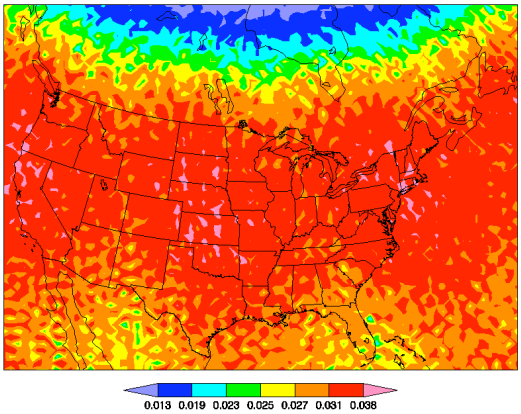


Figure 3.2.20: 1-Profile7D/Profile7T

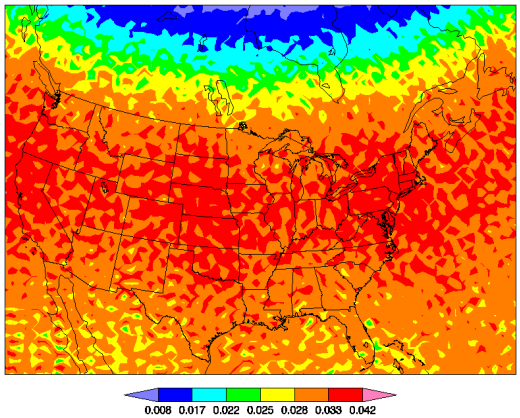


Figure 3.2.21: 1-Profile8D/Profile8T

Profiles 5 through 8 attempts to demonstrate the effect of DAK and TOMRAD on NO_2 at different pressure levels. These profiles were creating artificially and have the same integrated column concentration. Figures 3.2.14 through 3.2.17 show the results for the first day of July. The DAK port side bias decreases with increasing altitude. Figures 3.2.18 through 3.2.21 show the monthly average results. Likewise, the port side bias falls off with increasing altitude of NO_2. Nevertheless, the solar zenith angle bias becomes increasingly evident and maximizes at about 40 degrees latitude. Of the profiles tested, the maximum monthly average relative difference at altitude is only about 5%.

3.3 REAM vs. TM4 Model Differences

There are several differences between the models REAM and TM4. For the retrieval, model temperature and NO_2 mixing ratio profiles are used. The direct comparison between the two shows how and where to expect different air mass factor behavior. REAM produces cooler temperature profiles than TM4 across the majority of continental North America with a maximum relative difference of about 5%. Off the pacific coast REAM's temperature profile is warmer than TM4 with a maximum relative difference of about -2.5%. These differences are on the order of a few degrees Kelvin and represents only 1% of the calculated air mass factor error. REAM and TM4 are in good agreement with temperature.

At the near surface, REAM generally has higher NO_2 mixing ratios than TM4. Some exceptions include the states of New York, Pennsylvania, Missouri, California, and Nevada. These regions are about 50% lower. On the other hand, nearly most of the major metropolitan cities in the southeast and southwestern United States of America stand out with higher NO_2 mixing ratios at the surface of REAM, Figure 3.3.14. The higher resolution of REAM does capture the chemistry in close proximity to these cities. Some of these cities have as much as double the NO_2 represented in TM4. On Canadian surface, it is generally observed that REAM NO_2 is 50% to 100% higher than TM4. NO_2 mixing ratios are small here and the relative differences can be expected to be large. Atlantic Coast NO_2 is 50% to 100% less in REAM than in TM4, indicating that REAM does not move NO_2 as far as TM4.

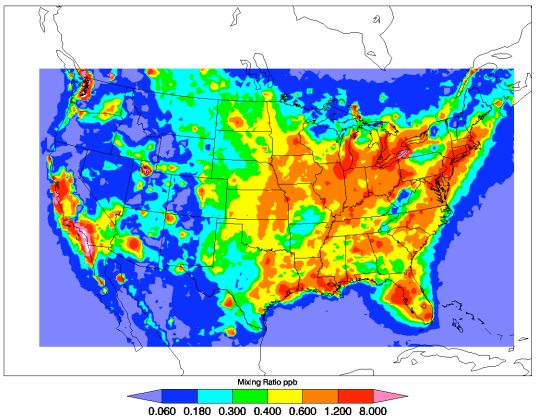


Figure 3.3.1: REAM36Ms

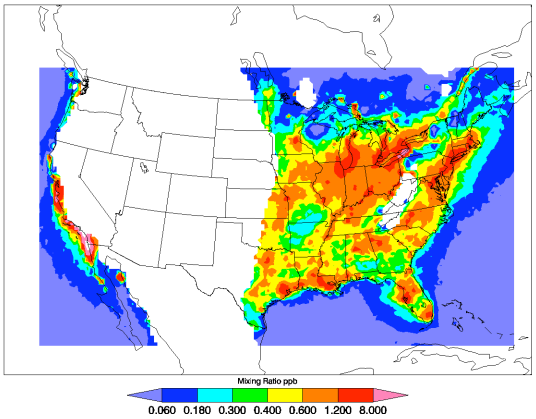


Figure 3.3.2: REAM36M8

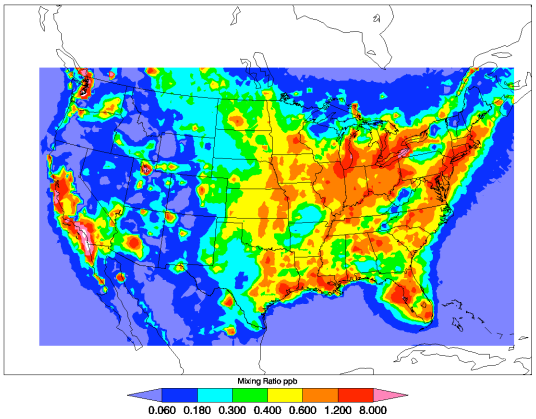


Figure 3.3.3: REAM36MnoLTs

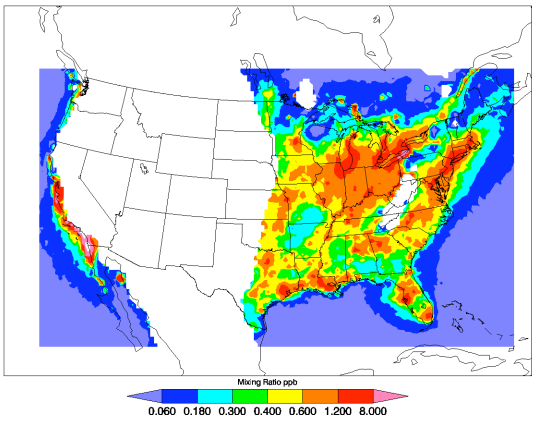


Figure 3.3.4: REAM36MnoLT8

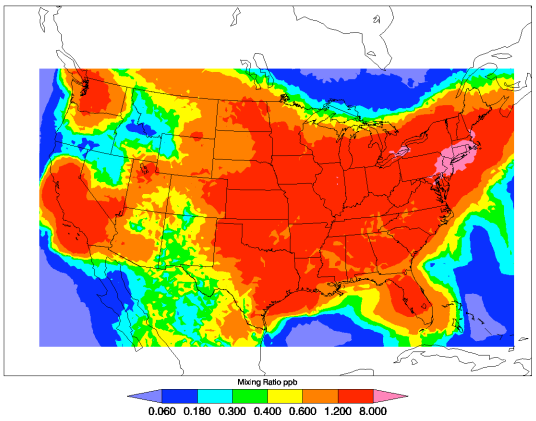


Figure 3.3.5: TM4Ms

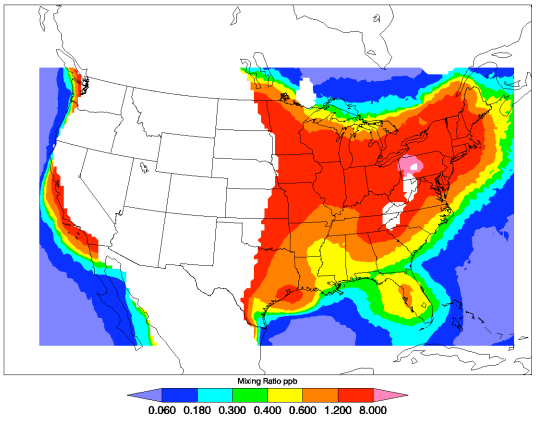


Figure 3.3.6: TM4M8

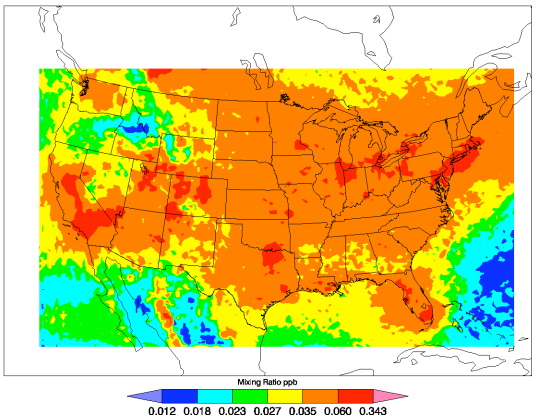


Figure 3.3.7: REAM36M5

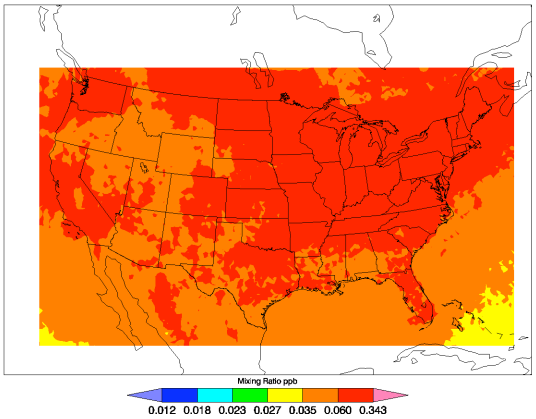


Figure 3.3.8: REAM36M3

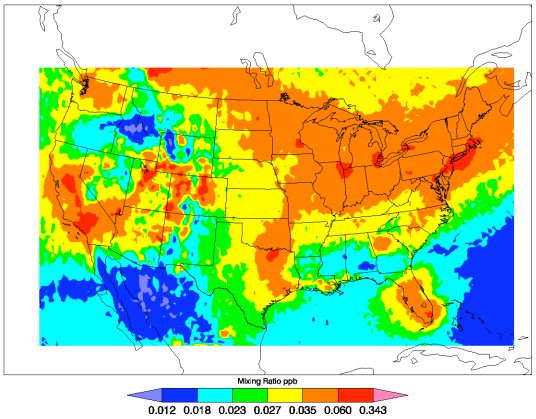


Figure 3.3.9: REAM36MnL5

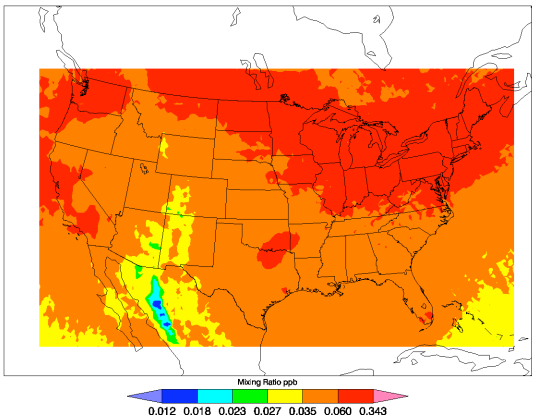


Figure 3.3.10: REAM36M3

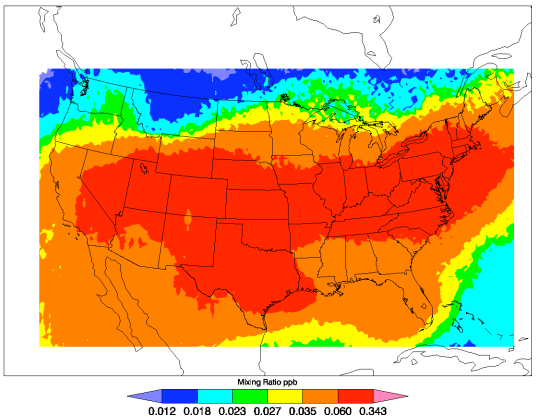


Figure 3.3.11: TM436M5

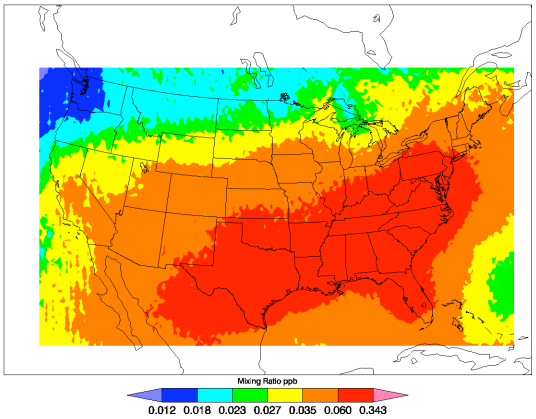


Figure 3.3.12: TM436M3

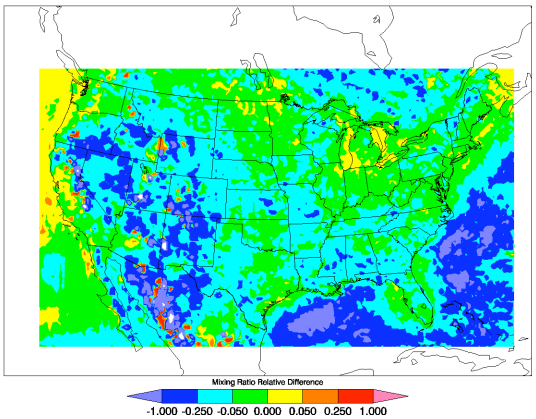


Figure 3.3.13: 1-REAM36Ms/REAM36MnLs

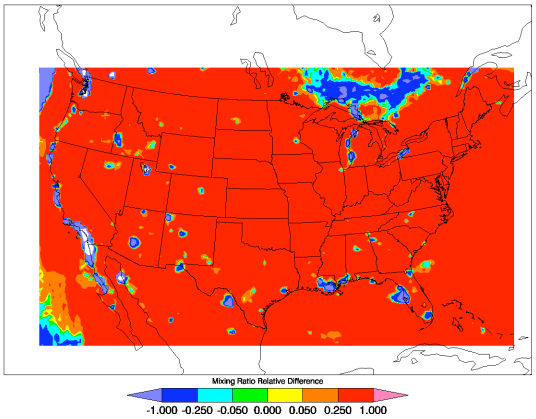


Figure 3.3.14: 1-REAM36Ms/TM436Ms

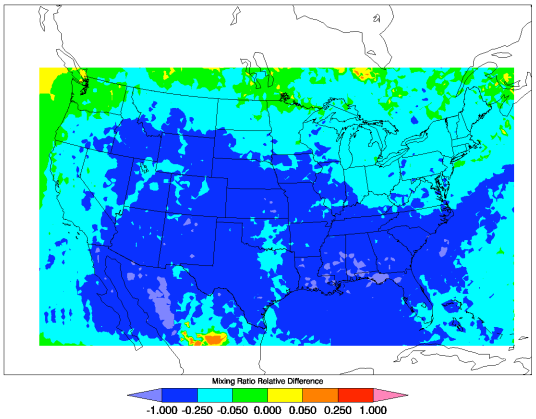


Figure 3.3.15: 1-REAM36M5/REAM36MnL5

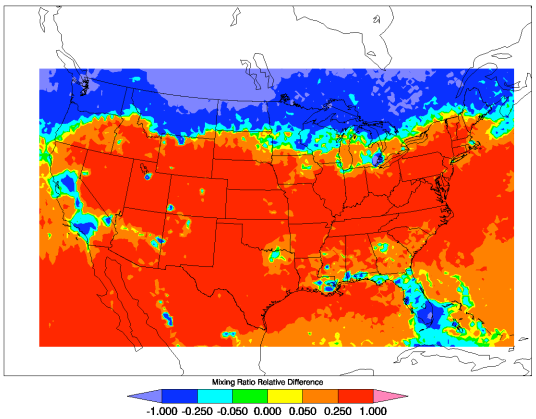


Figure 3.3.16: 1-REAM36M5/TM436M5

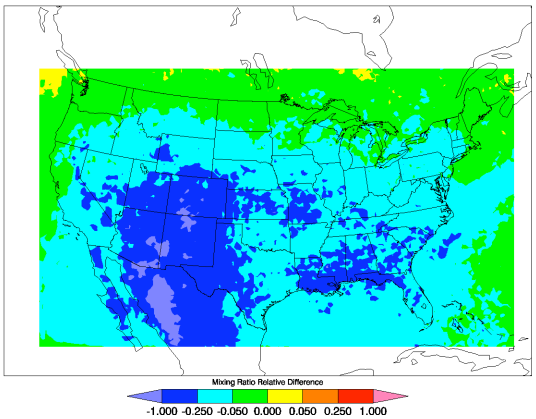


Figure 3.3.17: 1-REAM36M3/REAM36MnL3

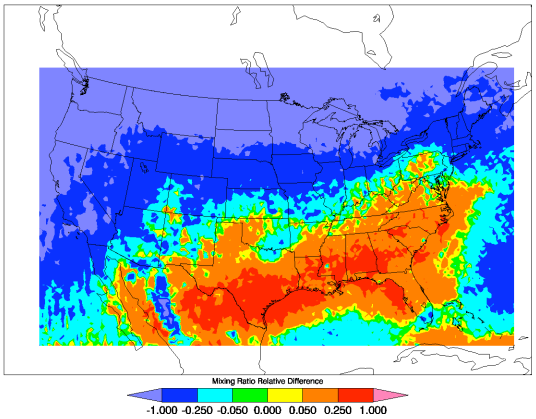


Figure 3.3.18: 1-REAM36M3/TM436M3

At 800 mb some of the same features may be observed as at the surface. Comparing REAM to TM4, these features included higher NO_2 mixing ratios above cities in the southeast and the broader Mississippi valley and lower NO_2 mixing ratios over much of the rural areas. All of the Atlantic off-shore regions are still 50–100% lower. Off the coast of British Columbia REAM is higher than TM4 likely due to bioactivity. Hudson Bay and the Labrador Sea also exhibit this behavior.

In the upper atmosphere, at 500 mb, most of the effects of cities are gone. At this level, Canadian NO_2 mixing ratios are 25% to 100% higher with REAM than those in TM4. This difference could be a result of a combination of things including better Canadian emissions inventories and thunderstorm activity. At the 200 mb level, lightening activity over Mexico and the southwest dominant the difference. A similar hot spot is found over central Florida and the Caribbean. Over Canada, it is observed that TM4 and REAM are exactly equal as it would be expected since the TM4 stratosphere is used with both models.

Figures 3.3.19 through 3.3.22 above show the vertically integrated tropospheric a priori models results, averaged for the month of July 2006 used in this comparison. Of the three models, TM4 has the lowest resolution and the least amount of NO_2 spatial variability. The REAM 36 km model, conversely, has the highest resolution and accounts better for urban and rural NO_2 variability. The REAM 70 km model covers a wider domain and with nearly a quarter the resolution of REAM 36 model, but represents a twofold improvement in resolution over TM4. The REAM models are better at

capturing urban and power plant NO_2 variability compared to TM4, with power production features like the Ohio River Valley and Four Corners. The REAM models also pay more attention to cities Portland, Oregon and Atlanta, Georgia. REAM 36km further acts to concentrate NO_2 to the source region compared to REAM70 and TM4.

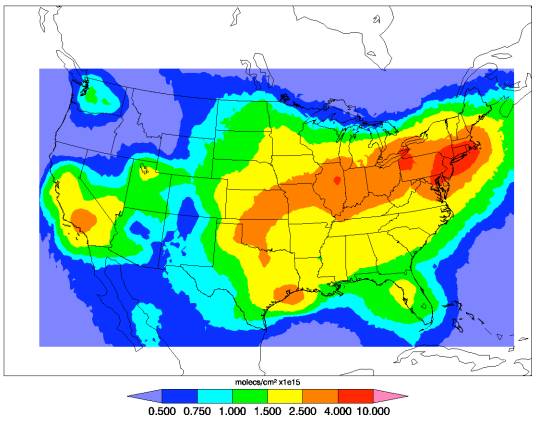


Figure 3.3.19: TM4M

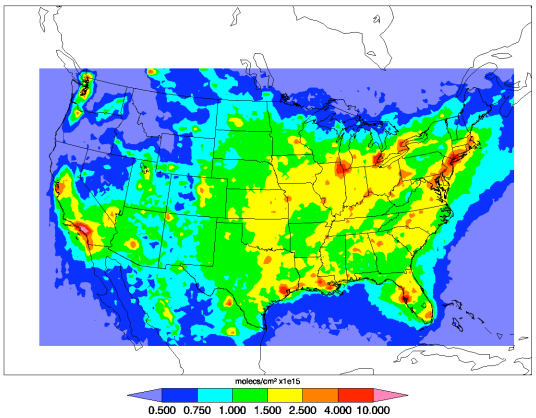


Figure 3.3.20: REAM36M

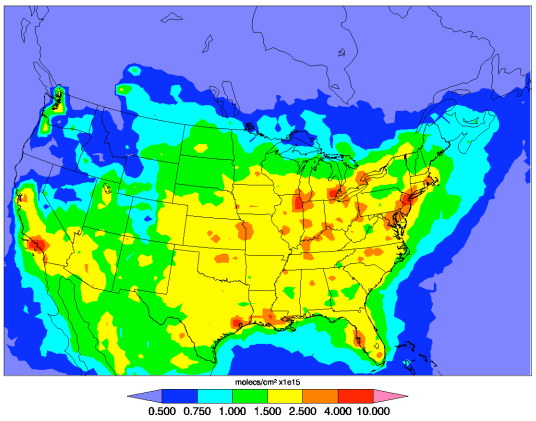


Figure 3.3.21: REAM70M

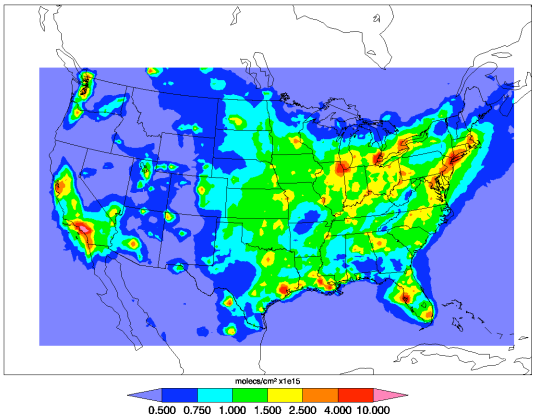


Figure 3.3.22: REAM36MnL

Figure 3.3.23 shows the relative difference between the REAM 36km and REAM70km models. A negative value indicates that REAM36 is higher and a positive value indicates that REAM70 is higher. In the rural areas and over much of the ocean, REAM 36 is 25 to 50 percent lower than REAM70. In the eastern half of Canada and the mountains of Mexico, REAM36 is more the 50% lower than REAM70. Most of the urban regions are more than 50 percent higher in the REAM36 model compared with the REAM70 model. There are a few off shore differences in the REAM36 near South Carolina and Southern California that do not appear in the direct model output vertical column.

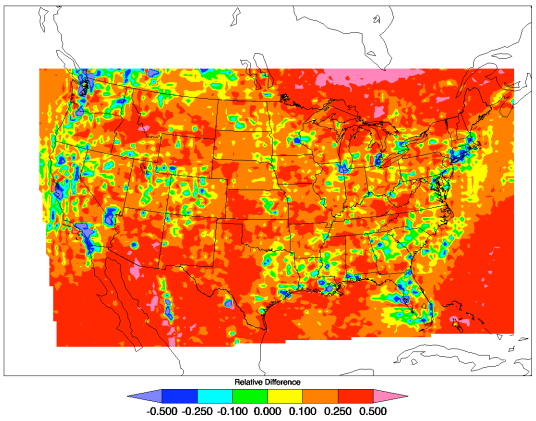


Figure 3.3.23: 1-REAM36M/REAM70M

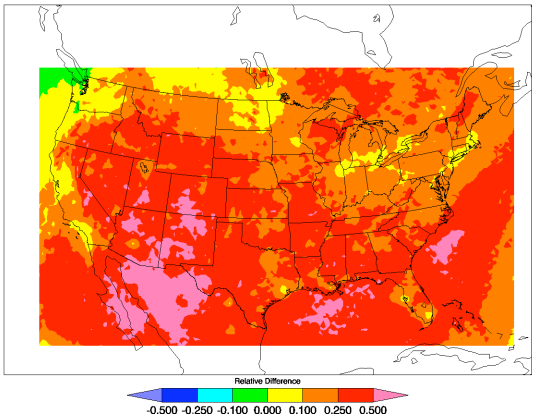


Figure 3.3.24: 1-REAM36MnL/REAM36M

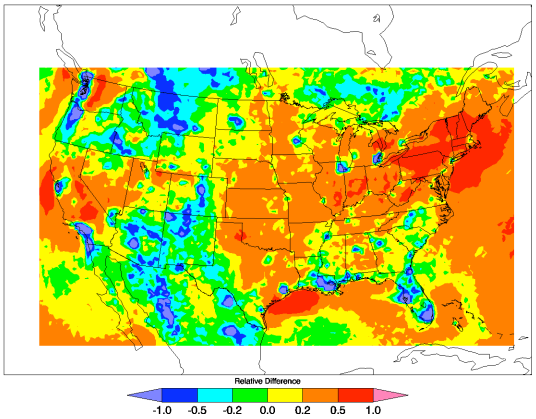


Figure 3.3.25: 1-REAM36M/TM436M

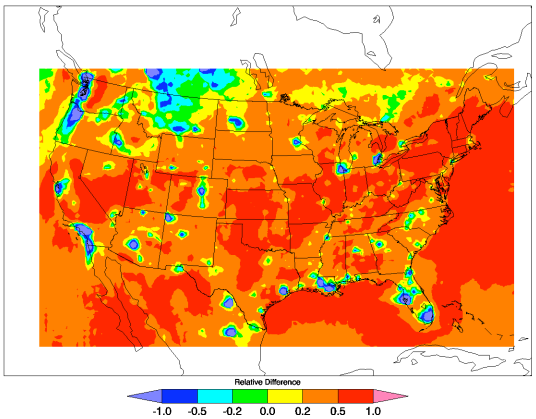


Figure 3.3.26: 1-REAM36MnL/TM436M

3.4 Retrieval Comparison

The purpose of this test is to show how the retrieval reacts to different modeled a priori profiles in addition to each radiative model. These methods were discussed in section 2.4. Figure 3.5.1 shows the results from this test. The following results indicate how these retrievals compare relatively to one another.

Figure 3.4.1 shows how REAM and TM4 compare over the 1st of July at the time of OMI overpass. The green bulleted series indicate the model vertical column. The violet bulleted series indicate the KNMI retrieval slant column calculated using that model vertical column. The plot serves multiple purposes. First, it is to show how REAM vertical structure compares with TM4 vertical structure. They are nearly identical in structure in number of vertical layers. It also shows how REAM merges into the TM4 stratosphere. This process left few discontinuities. And finally, the plot shows show how the retrieval is based on the distribution of NO_2 in the vertical column and not the

magnitude of the vertical column. REAM has relatively high slant columns aloft due to sensitivity to the radiation model here. A higher slant column means a higher air mass factor which means a lower retrieval vertical column. In the case of downtown Atlanta, REAM's small slant column near the surface reverses this effect.

Figure 3.4.2 indicates that REAM has at times significantly more NO_2 in the upper troposphere compared with TM4. In the upper troposphere the lifetime of NO_2 is on the order of days. Such long lasting events tend to reduce REAM retrieval vertical columns as result of the air mass factor calculation. Air mass factors are a function of the a priori distribution which is then scaled by the radiative transfer model. These events spread out to have much as half the REAM domain. Unlike in the urban case, Figure 3.4.1, there low near-surface NO_2 to make up for the difference.

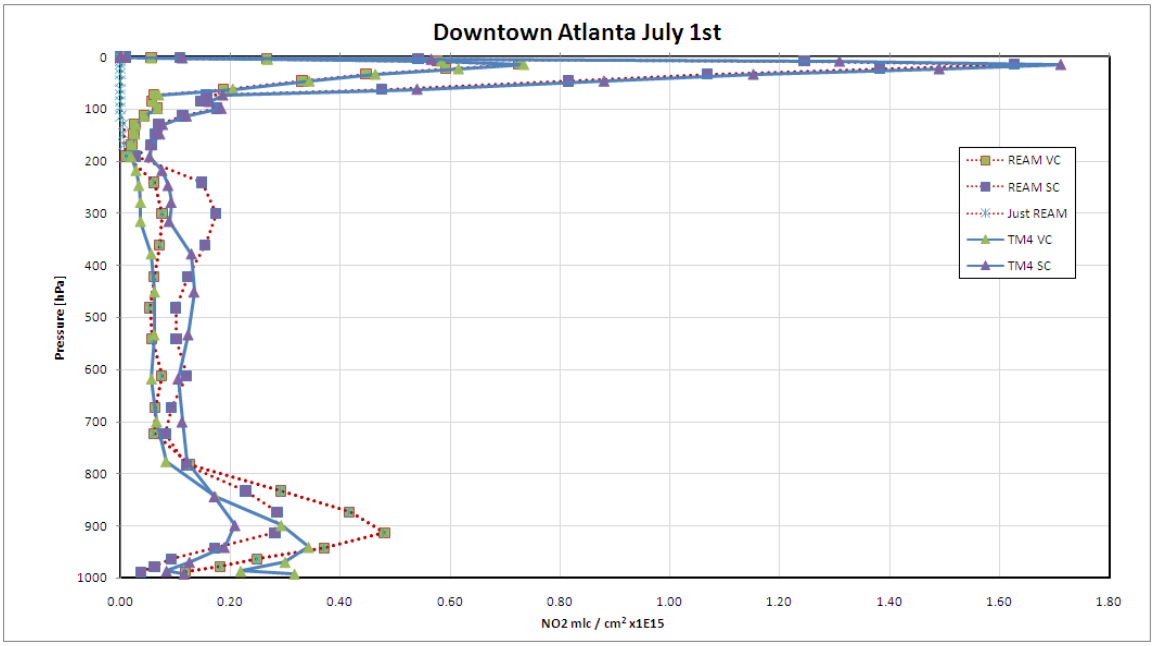


Figure 3.4.1: Comparison of the Model Vertical and Retrieval Slant Column, Atlanta

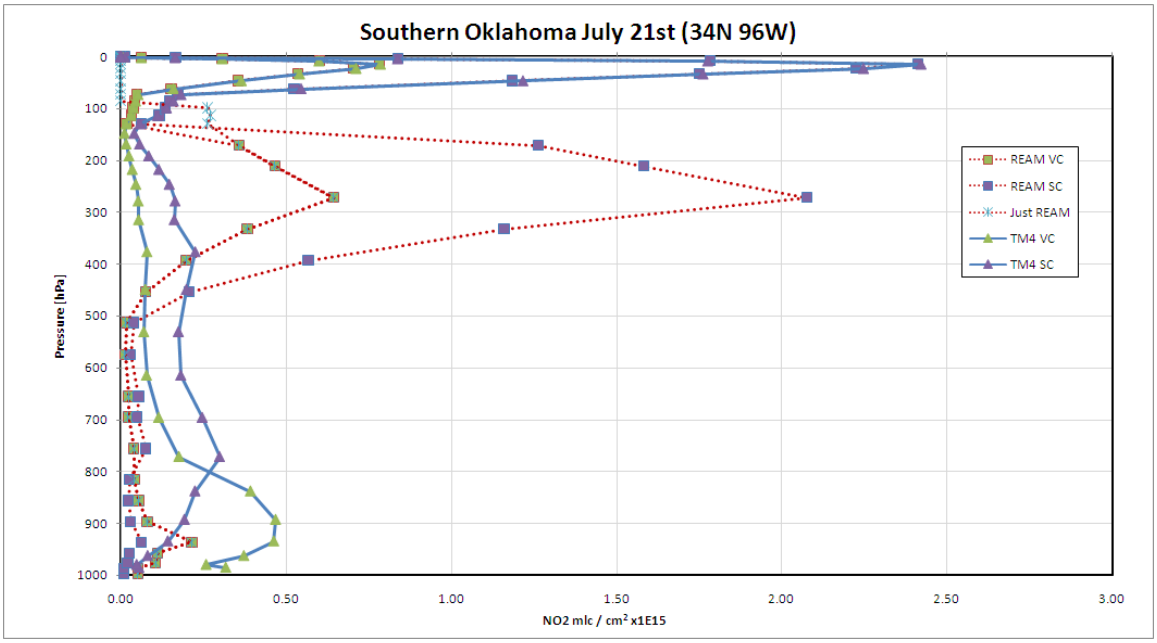


Figure 3.4.2: Comparison of the Model Vertical and Retrieval Slant Column, Oklahoma.

Figures 3.4.3 through 3.4.8 show the direct retrieval results for the three a priori models above using the KNMI NO_2 radiance model DAK and TOMRAD averaged over the month of July. The general features of the retrieval somewhat mimic their respective a priori NO_2 models, as shown in Figure 3.3.19 through 3.3.21. The retrieval thus has a bias to the chemical transport model which initializes it. The retrieval does plenty, however, to bring the three different a priori models, TM4, REAM70, REAM36, closer to the same result given the same set of satellite observations. The direct relative differences between the three retrievals are described later.

REAM70 retrieval is closer to resembling TM4 retrieval than the REAM36 retrieval. REAM36 shows a significantly cleaner Rocky Mountains region. In addition, the urban centers are much better defined and the coastal NO_2 is much reduced as seen in the REAM36.

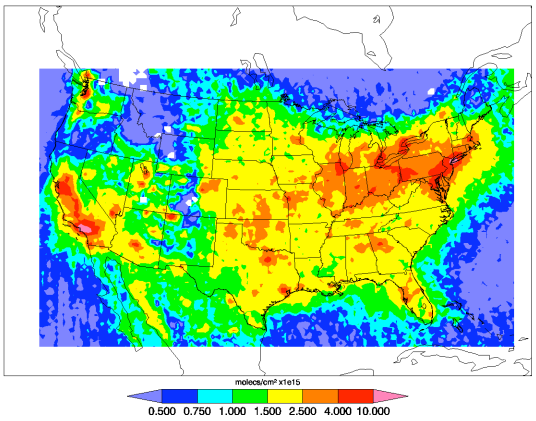


Figure 3.4.3: TM4D

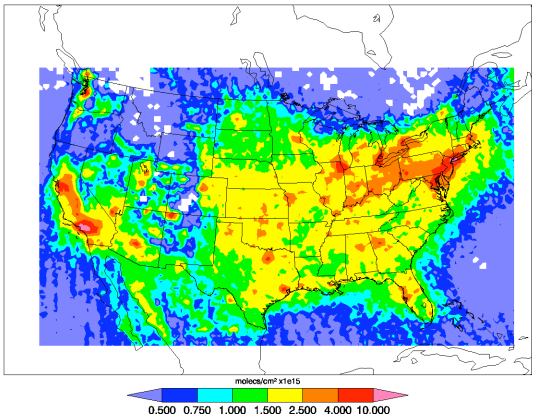


Figure 3.4.4: TM4T

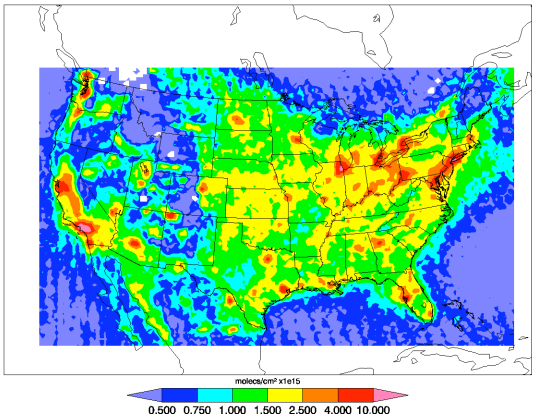


Figure 3.4.5: REAM36D

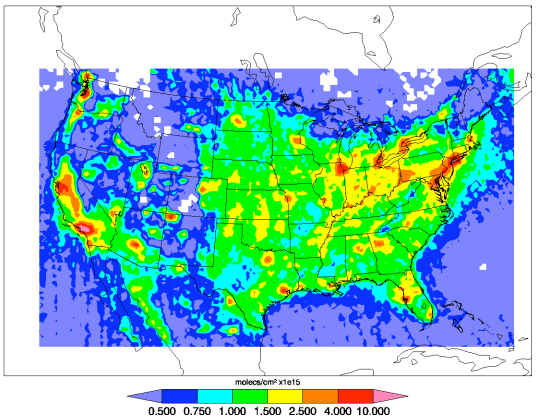


Figure 3.4.6: REAM36T

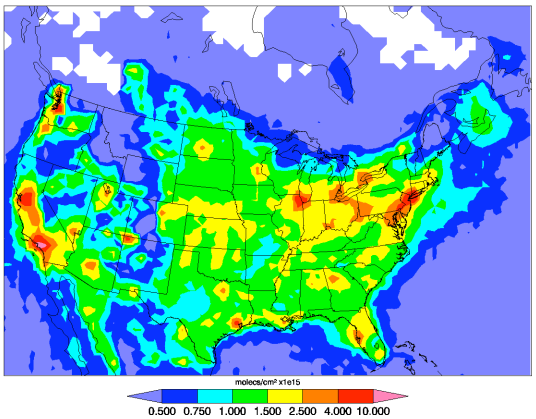


Figure 3.4.7: REAM70D

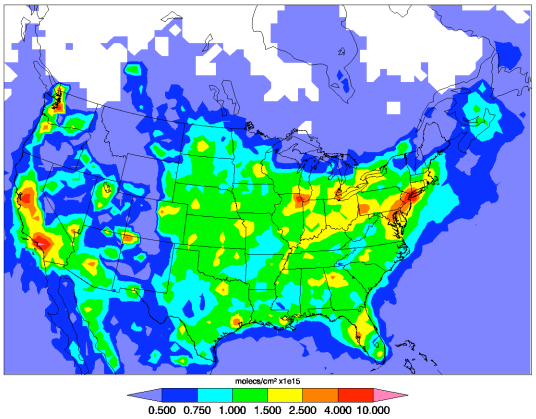


Figure 3.4.8: REAM70T

Figure 3.4.9 represents the relative difference between the retrieval products using REAM36 and REAM70 a priori models. It is a busy plot due to the resolution interpolation scheme used, but some general patterns emerge. A negative value indicates that REAM 36 retrieval produces a larger result. Across the board, REAM36 a priori retrieval shows more NO_2 in urban and industrial regions. However, in the high elevation rural regions of the western United States there is remarkably less NO_2 in the REAM36 retrieval which is in contrast to the lowland Mississippi River states rural regions. A similar pattern was not found in the direct model comparison, meaning that the difference in east and west is as result of the DAK radiation model.

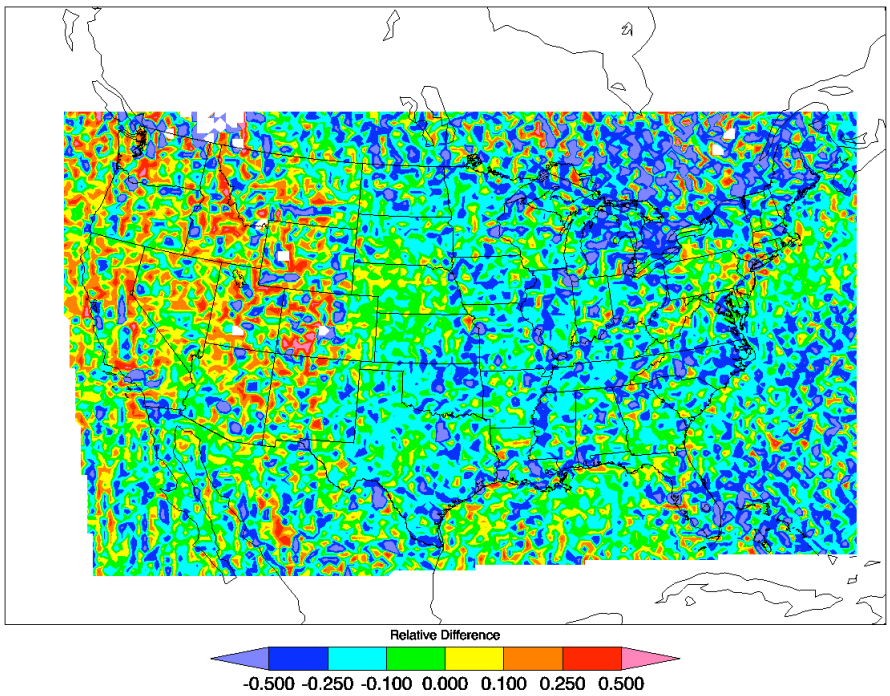


Figure 3.4.9: Relative difference 1-REAM36/REAM70
NO_2 Vertically Integrated Retrieval with DAK radiation

Figures 3.4.10 through 3.4.13 compare four different versions of the TM4 model to the TM4 retrieval. Red colors indicate that the model is too low. Figures 3.4.10 and 3.4.12 use KNMI's DAK radiation model and Figures 3.4.11 and 3.4.13 use NASA's radiation model TOMRAD. Figures 3.4.10 and 3.4.11 use the REAM36 grid and Figures 3.4.12 and 3.4.13 use the REAM70 grid. When comparing the model and retrieval, grid size matters. For example, Figure 3.4.12 shows less extreme differences between the TM4 model and the DAK retrieval compared with Figure 3.4.10. The same pattern due to grid resolution alone is observed using TM4 together with TOMRAD.

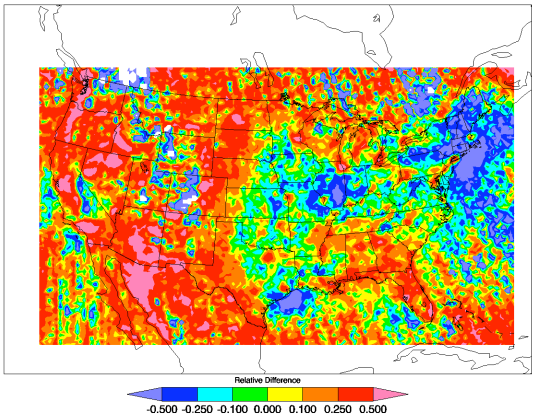


Figure 3.4.10: 1-TM436M/TM436D

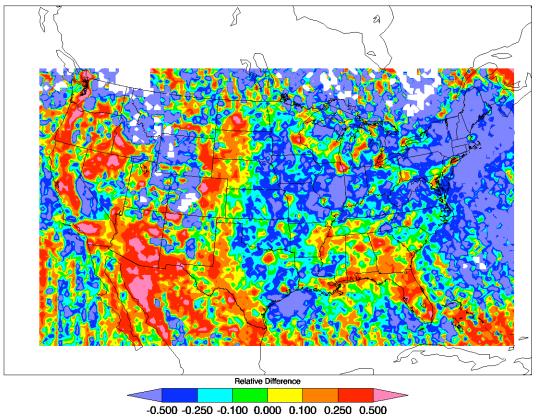


Figure 3.4.11: 1-TM436M/TM4T

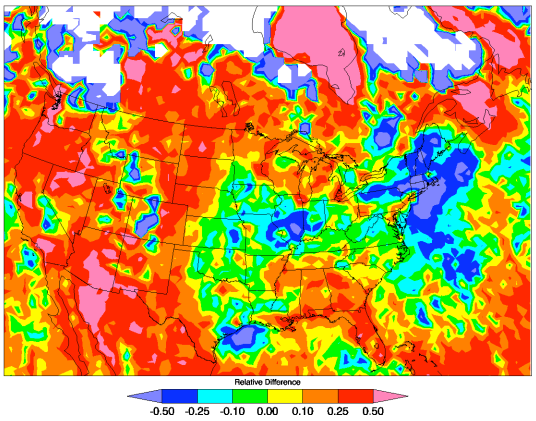


Figure 3.4.12: 1-TM4M/TM470D

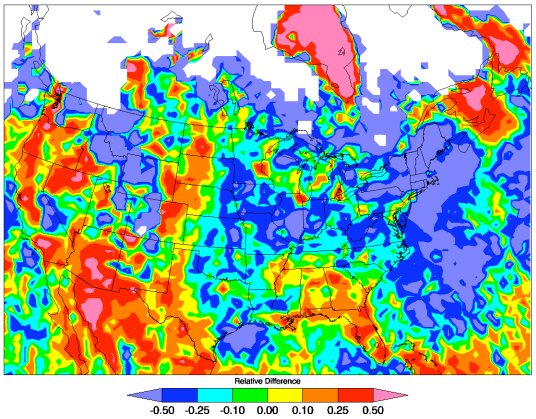


Figure 3.4.13: 1-TM4M/TM470T

The TM4 model with the DAK retrieval compares better than the TM4 model with TOMRAD over the Ohio Valley and the plains states west of the Mississippi River, Figures 3.4.10 and 3.4.11. However, the TM4 model underestimates much of the Rocky Mountain states and the Southeast. However, TOMRAD shows improvement in the Southeast compared with the DAK retrieval and exhibits a more balanced positive and negative difference over the Rocky Mountain States. Both DAK and TOMRAD retrievals indicate that TM4 underestimates a priori profiles for most urban and industrial regions except for the off-shore Texas oil rigs.

Figures 3.4.14 through 3.4.17 show how the REAM a priori model compares with the DAK and TOMRAD retrievals. Unlike in the last comparison, the REAM70 and REAM36 are models that are simulated separately. The resulting model differences are shown in Figure 3.4.9. The scale on Figures 3.4.16 and 3.4.17, the REAM70 simulation, are doubled. The REAM36 model offers superior a priori profiles compared to the REAM70 model using either radiation model. The TOMRAD retrieval is less sensitive to REAM36 a priori model and comes to better agreement in urban regions, suggesting that the REAM a priori overestimates rural regions with the exception of the west coasts of the United States and Mexico. The DAK model, however, suggests that REAM36 underestimates border regions of the west coast, the North Midwest and Great Lake states. Thus, depending on the question, one radiative model offers better agreement then the other.

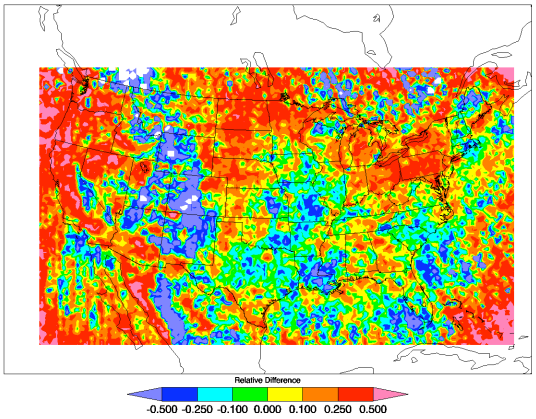


Figure 3.4.14: 1-REAM36M/REAM36D

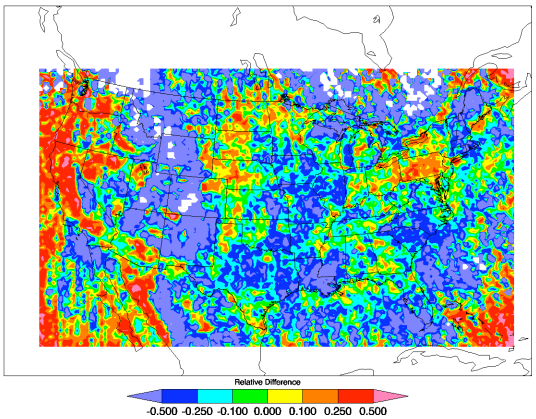


Figure 3.4.15: 1-REAM36M/REAM36T

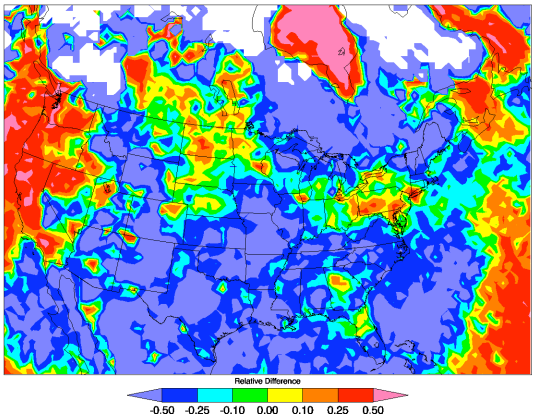


Figure 3.4.16: 1-REAM70M/REAM70D

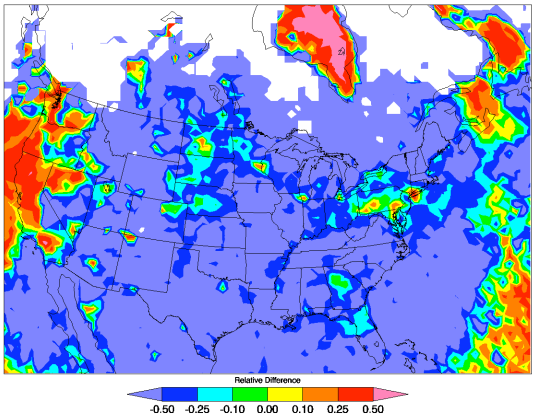


Figure 3.4.17: 1-REAM70M/REAM70T

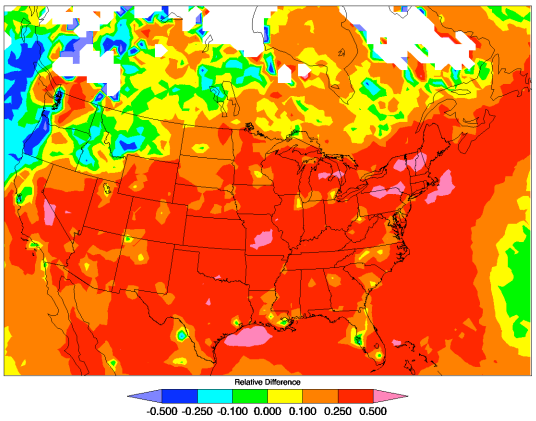


Figure 3.4.18: 1-REAM70D/TM470D

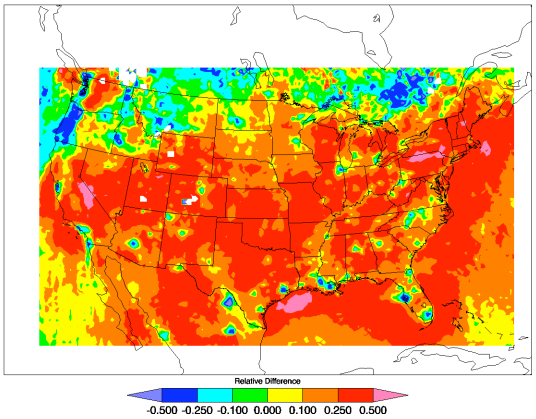


Figure 3.4.19: 1-REAM36D/TM436D

In general REAM36 and REAM70 DAK retrievals compare similarly with the TM4 DAK retrieval, Figures 3.4.18 and 3.4.19. The result consists of much lower greater than 25% NO_2 over the majority of the United States. The exception is the Pacific Northwest where REAM and TM4 DAK retrievals agree more closely. On the Oregon coast line, the REAM DAK retrieval produces more NO_2 than TM4. The largest difference between REAM a priori profiles and TM4 a priori profiles are found in urban centers of the United States. The exception being the New England states where TM4 profiles have their highest boundary layer concentration over this domain. In general, REAM a priori profile retrievals put more NO_2 in urban and industrial regions and less NO_2 in rural regions.

The amount of lightning also impacts REAM retrievals. Removing the lightning increases the retrieval NO_2 because the model a priori shifts the NO_2 distribution toward the boundary layer where the radiation model is more sensitive. DAK is more sensitive than TOMRAD. As discussed in the next section, DAK also has a stronger port side bias then TOMRAD. Hence, TOMRAD much more balanced.

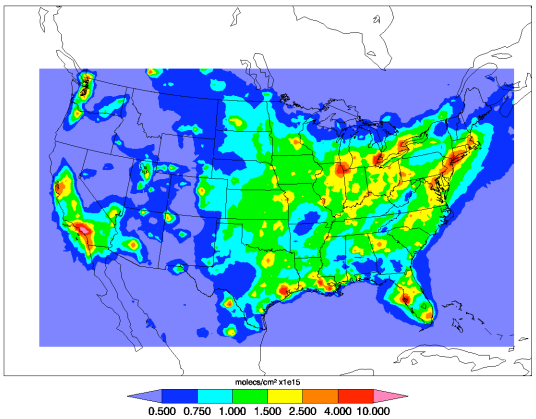


Figure 3.4.20: REAM36MnL

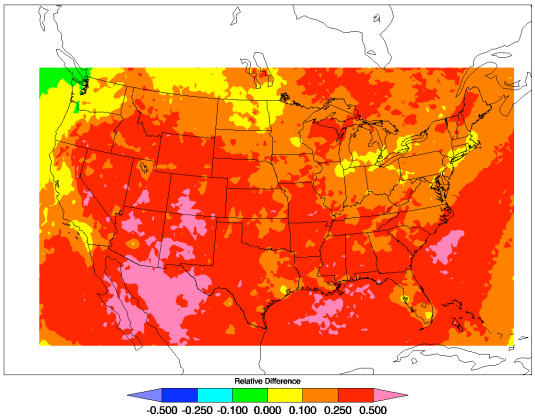


Figure 3.4.21: 1-REAM36MnL/REAM36M

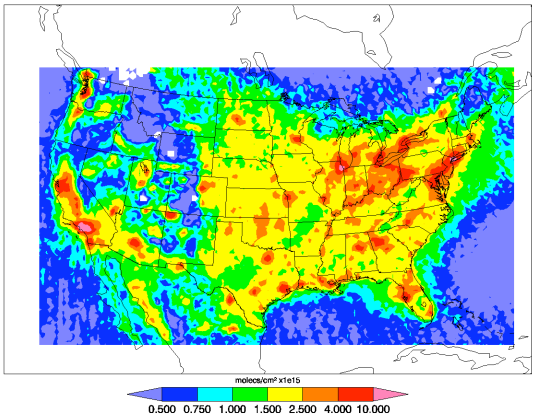


Figure 3.4.22: REAM36DnL

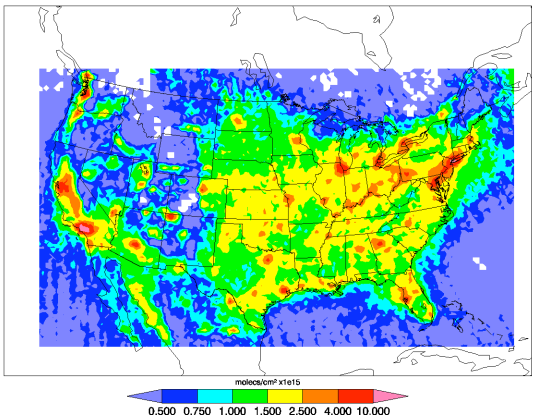


Figure 3.4.23: REAM36TnL

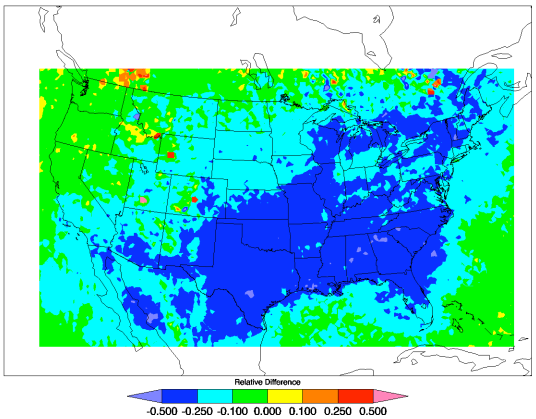


Figure 3.4.24: 1-REAM36DnL/REAM36D

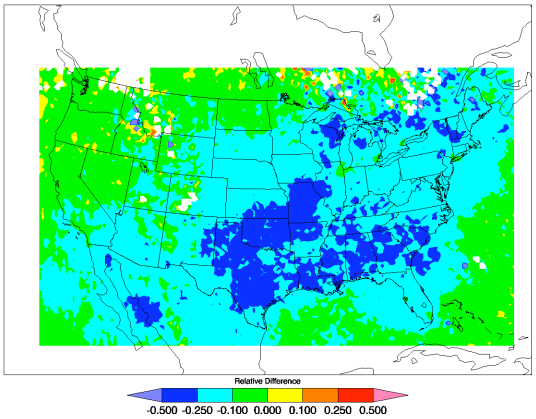


Figure 3.4.25: 1-REAM36TnL/REAM36T

3.5 Internal Retrieval Issues

The retrieval itself has issues that confound these results. As discussed in greater detail in the section 1.1, because of the geometry of the OMI satellite's sensor, pixels observed port and starboard are not viewed symmetrically. The viewing zenith angle of the satellite over a certain coordinate on the ground flips from high one day (~40-60°) to low (~0°) the following day, repeating. The result is that that coordinate is viewed daily from an alternating satellite port and starboard side perspective, generating an alternating retrieval bias for that coordinate. In addition to this alternation, the satellite overpass path shifts eastward of the coordinate with a period of about eight days before the path is directly overhead again. The NO_2 radiative transfer model attempts to make up for this asymmetry as a function of relative azimuth angle (RAA). The results in this section show to what extent this correction is an issue for the retrieval, and focuses on the DAK NO_2 radiation model. TOMRAD has a less extensive but similar issue.

Figures 3.5.1 through 3.5.6, show two sets of images, over two different period of days. Figures 3.5.1, 3.5.3, and 3.5.5 show a period of 8 days starting on July 8th and Figures 3.5.2, 3.5.4, and 3.5.6 show a period of 30 days starting on July 1st. The periods of days are broken down into even and odd days where even days are averaged together and odd days are likewise averaged together. The subsequent image is the relative comparison of these two averages, odd days and even days.

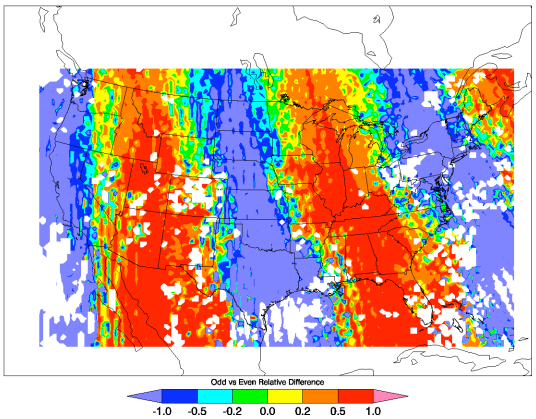


Figure 3.5.1: 1-RAAodd/RAAeven 5-8

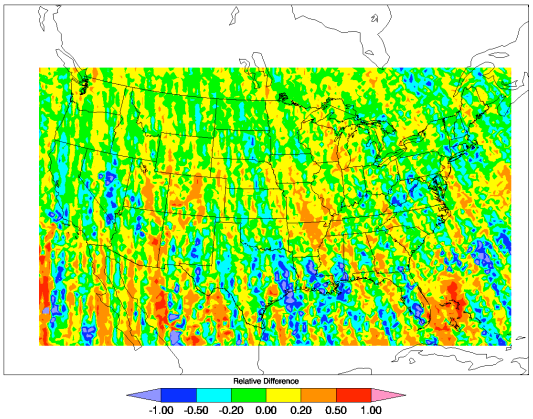


Figure 3.5.2: 1-RAAodd/RAAeven 1-30

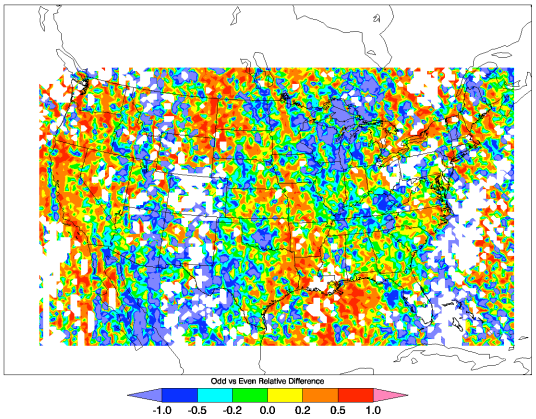
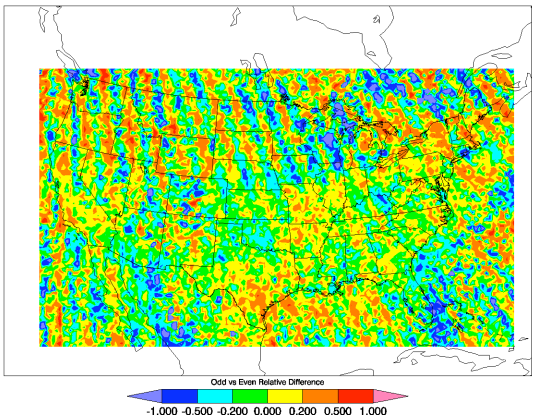


Figure 3.5.3: 1-REAM36Do/REAM36De5-8 Figure



3.5.4: 1-REAM36Do/REAM36De1-30

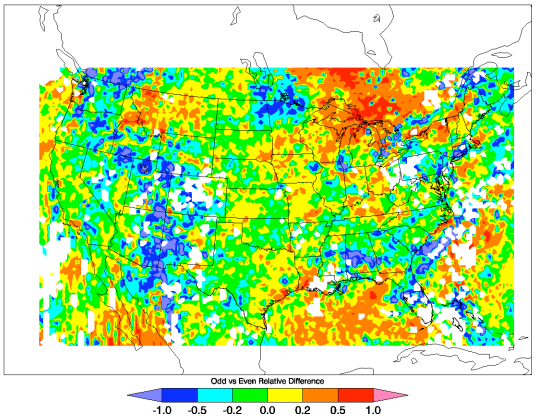


Figure 3.5.5: 1-REAM36Mo/REAM36Me5-8

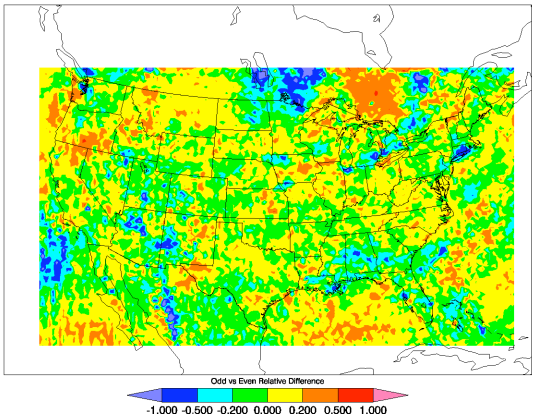


Figure 3.5.6: 1-REAM36Mo/REAM36Me1-30

Figures 3.5.1 and 3.5.2 show the RAA odd and even day comparison over each respective period of days. Figures 3.5.3 and 3.5.4 show the retrieval corresponding with the figures directly above. Figures 3.5.5 and 3.5.6 show the same averaging technique applied to the REAM36 Model. It is evident, looking at the retrieval and the RAA that there is a strong anti-correlation between the two. This indicates an odd day, even day bias which is not observed in the REAM36 model. For a particular coordinate, it is not evident how what is NO_2 variability or RAA radiative model over correction. Comparing odd and even days over a whole month helps to remove this bias. However, there are many "slits" due to the eastward swath progression that escape the average. Thus, if a particular coordinate of interest falls in one of these slits and day-to-day variability is of concern, much of the variability may be artificial.

Figures 3.5.7 through 3.5.18 show the same sort of comparison as the previous figures. Nevertheless, in addition to contrasting odd and even days, the comparison is broken down by sets of every other day (same as comparing odd and even), every other set of 2 days, every set of 3 days, 4 days, 6 days and every other set 8 days over a period of 24 days starting on July 1st. The exception is that the 8 day comparison only looks at a 16 day interval. These are denoted in the figure titles as RAA1, RAA2, etc. and REAM36D1, REAM36D2, etc.

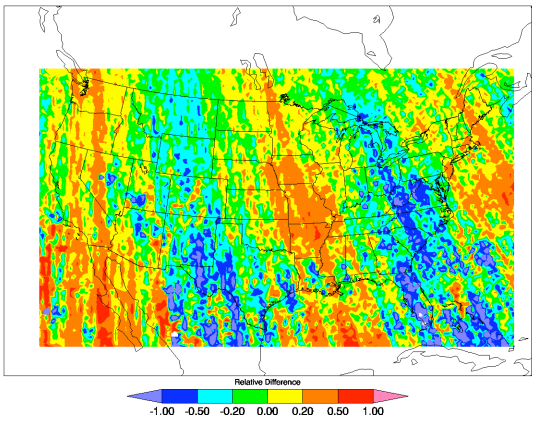


Figure 3.5.7: RAA1

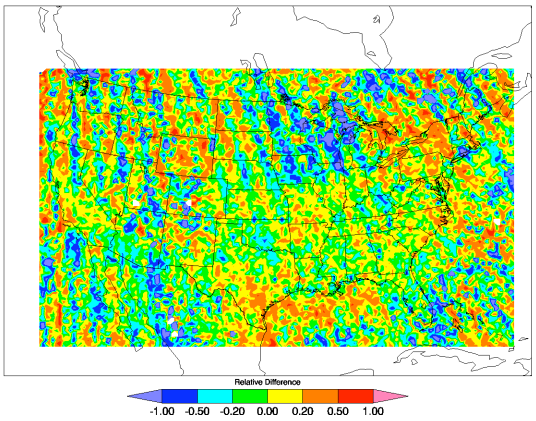


Figure 3.5.8: REAM36D1

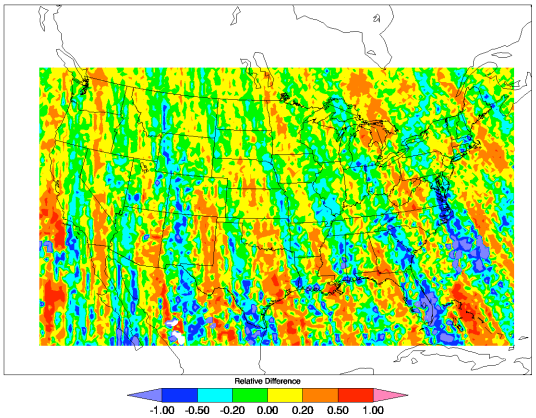


Figure 3.5.9: RAA2

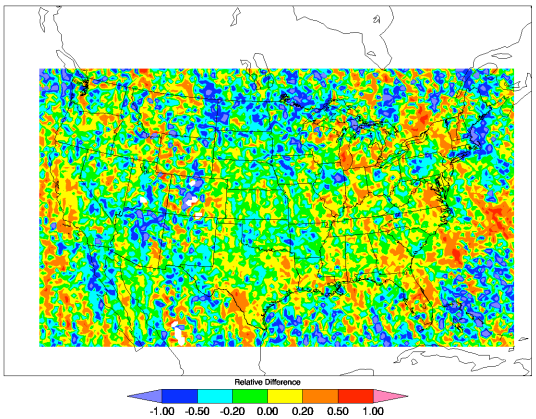


Figure 3.5.10: REAM36D2

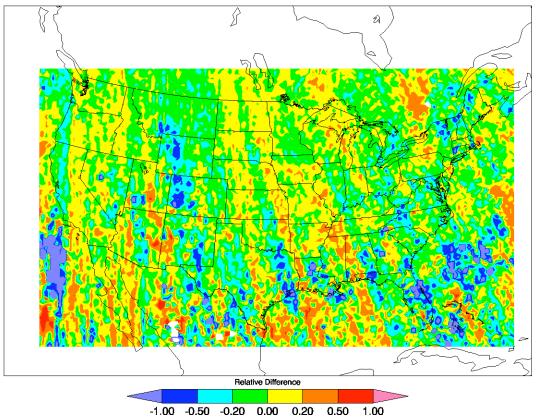


Figure 3.5.11: RAA3

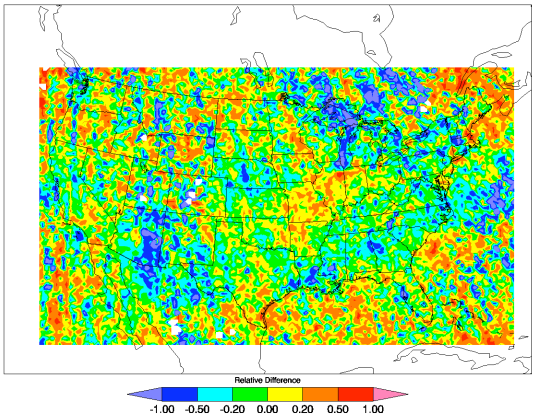


Figure 3.5.12: REAM36D3

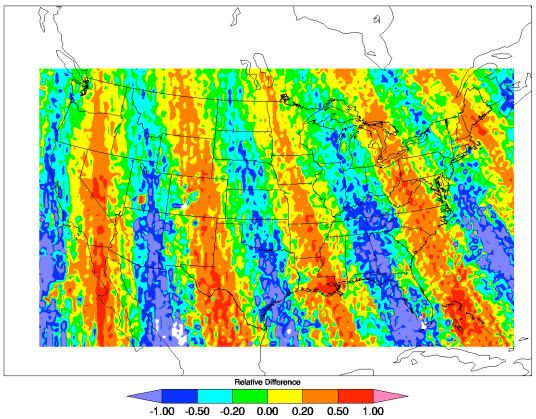


Figure 3.5.13: RAA4

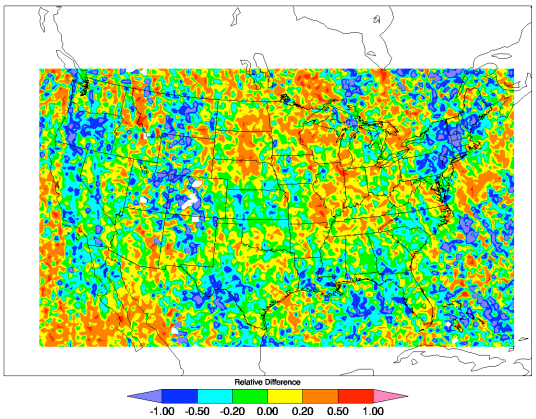


Figure 3.5.14: REAM36DAK4

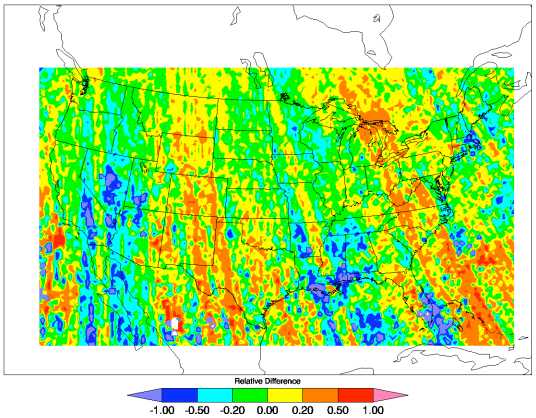


Figure 3.5.15: RAA6

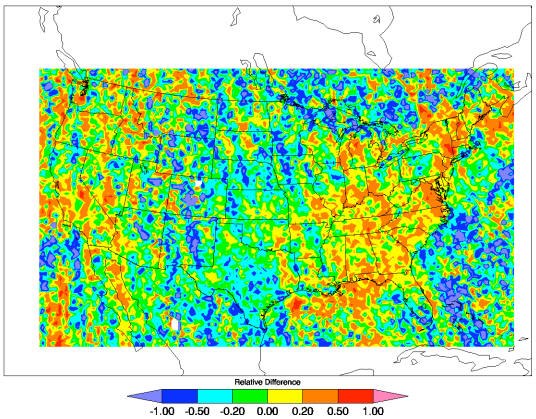


Figure 3.5.16: REAM36DAK6

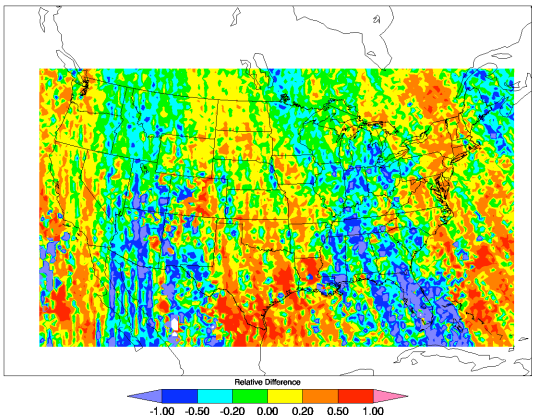


Figure 3.5.17: RAA8

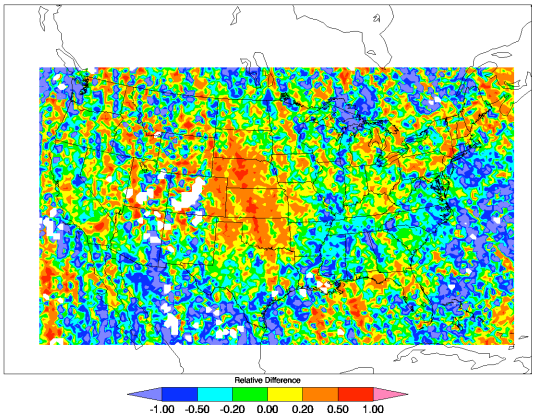


Figure 3.5.18: REAM36DAK8

Figures 3.5.7, 3.5.9, 3.5.11, 3.5.13, 3.5.15 and 3.5.17 show the RAA for 1 day, 2 day, 3 day, 4 day, 6 day and 8 day sets, respectively. To the right of these figures are the respective REAM36DAK retrievals. The reason that Figure 3.5.7 does not look like Figure 3.5.2 is that the retrievals are performed over a different period of days. Figure 3.5.7 is the first 24 days of July and Figure 3.5.2 is the first 30 days. Thirty days is closer to two complete sets of 8 day precession cycles whereas 24 days is more like 1.5 sets of 8 days precession cycles, hence Figure 3.5.7 shows a pattern that looks similar to figure 3.5.1.

The more centered on zero and the less correlated with RAA the retrieval the better. Thus, the three day average appears to do best. Average over every other day, every four days and every eight days do not perform well to eliminate the port side, starboard side observation bias due the daily alternating of viewing zenith angle and the eight day precession of the swath. In general, as the swaths become wider, nearer and nearer to the equator, the internal radiative model complication becomes worse.

3.6 Odd vs. Even Day Statistical Results on Retrievals for Select Cities

As observed in the previous results section, the satellite switches relative azimuth angle over a city each day. For example, one day the relative azimuth angle over Chicago will be less than 90 degrees, the next day it will be greater than 90 degrees, then the day after that it will be back to less than 90 degrees, and so forth.

Four cities were selected in North America to test for similarity of the REAM model to OMI satellite retrieval. These four cities include Atlanta, Los Angles, Chicago, and New York. Figure A.11 shows how OMI relates to the cities of interest over North America with respect to the viewing zenith angle. A viewing zenith angle of zero, blue, indicates the satellite is viewing straight down. In other words, at this location, the satellite passes over North America. The red indicates the maximum viewing zenith angle (~60°). This red zone, where viewing zenith angle is high, separates one swath from another. A little over one hour will elapse from one swath to the next. With reference to Figure A.11, port, or west, of the satellite is indicated by low relative azimuth angle and starboard, or east, of the satellite is indicated by high relative azimuth angle. Observe how New York is only port of the satellite on even days and starboard of the satellite on odd days. Similar relationships can be concluded with the other cities. This graphic is included to help in associating the satellite's viewing geometry in relationship to four cities for the following statistical analysis. As the days progress, the alternation of port and starboard move eastward.

To test the reality of these using real cities, a series of the figures below describe the results. The figures consist of two sections. The top section shows log normal fits of each of the 12 distributions: four different retrievals and the two chemical transport models used in this study, separated into even and odd days. Appendix A consists of the raw data, f-tests and t-tests, and the accompanying histograms for all the tests used in this presentation. Three tables which aim to perform statistical tests comparing one distribution to the next. The table on the bottom right shows how each of the 6 distributions compares on even days and odd days.

Generally speaking, both REAM and TM4 models alone are much more normally distributed than the retrieval. This implies that both models haven't yet captured much of the variability of day-to-day NO_2 column distributions.

Cities are evaluated from a start-day averaged for a period of days. The flexibility of start-day and range of days allowed one to consider the effects of satellite precession over a city depending on its location. Satellite coverage of a city is such that ever day VZA switches. If one day the city is nadir to the satellite the next day VZA will be 60 degrees. This switching back and forth processes eastward from one satellite swath to the next. It takes about 14-16 days before a city which observes nadir, observes nadir again. This implies a period of 8 days, 4 of which are observed port of the satellite and 4 of which are observed starboard of the satellite. From the constant profile test, section 3.3, it is expected that KNMI observations port of the satellite will yield larger NO_2 vertical columns.

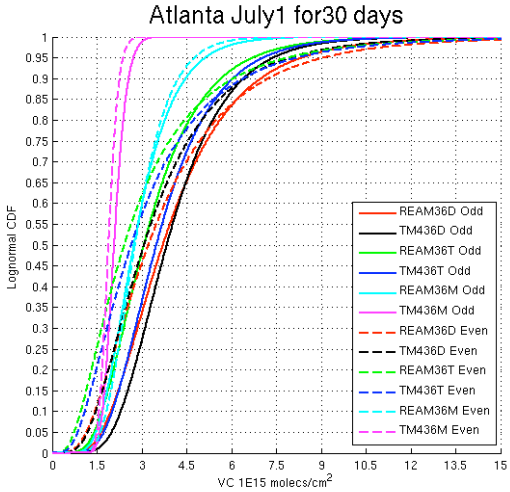


Figure 3.6.1: Atlanta 36 July1-30

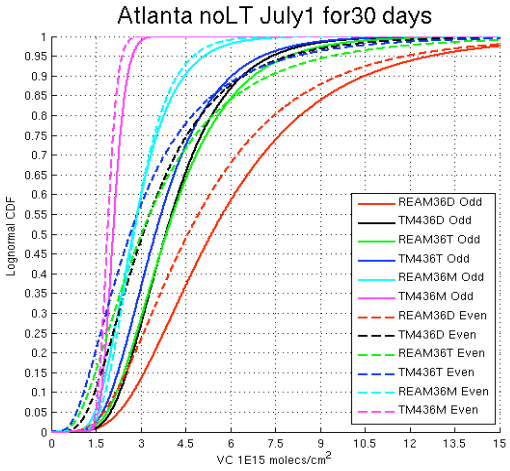


Figure 3.6.2: Atlanta 36nL July1-30

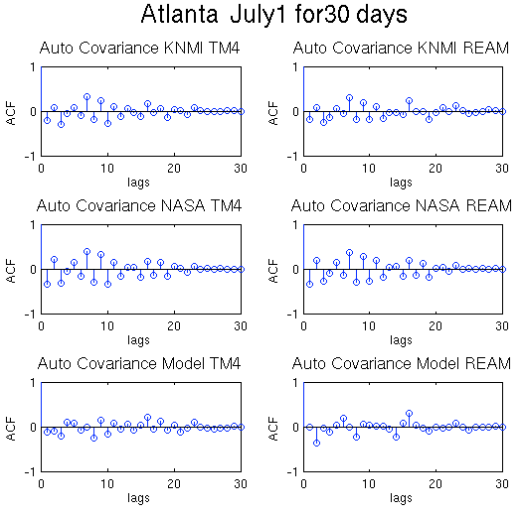


Figure 3.6.3: Auto covariance over Atlanta July1-30

Figures 3.6.1 and 3.6.2 show the analysis results for downtown Atlanta using REAM36 domain. The dashed lines are even days averaged together and the solid lines are the odd days averaged together. The red and black lines use DAK radiative transfer model and the green and dark blue color lines use the TOMRAD model. The magenta and light blue lines are the TM4 and REAM models respectfully. Ideally, the closer the lines are to one another, the closer the match of the various runs. Figure 3.6.2 is the 3.6.1 case except REAM lightning NOx production was turned off. The DAK retrievals are more sensitive to lightning NOx than the TOMRAD retrievals. By turning off lightning NOx, the retrieval to model compares worse than in the case with lightning NOx. Furthermore, the port side bias is clearer to observe in the no lightning NOx case.

Figure 3.6.3 shows the auto covariance of the each radiation and a priori model. The left column is dedicated to the TM4 model and the right column is dedicated to the REAM model. The top row uses the DAK model, the middle row uses the TOMRAD model, and the bottom row is the respective a priori model. Each lag indicates a subsequent day correlated with the first day of July. Odd days and even days are anti-collated, due to the port side bias as result of the flipping relative azimuth angle. Furthermore, the eight day satellite precession period is visible. However, these patterns are not evident in the a priori models. In addition, although the TOMRAD auto covariance is more balanced than DAK auto covariance the TOMRAD auto covariance is more strongly correlated.

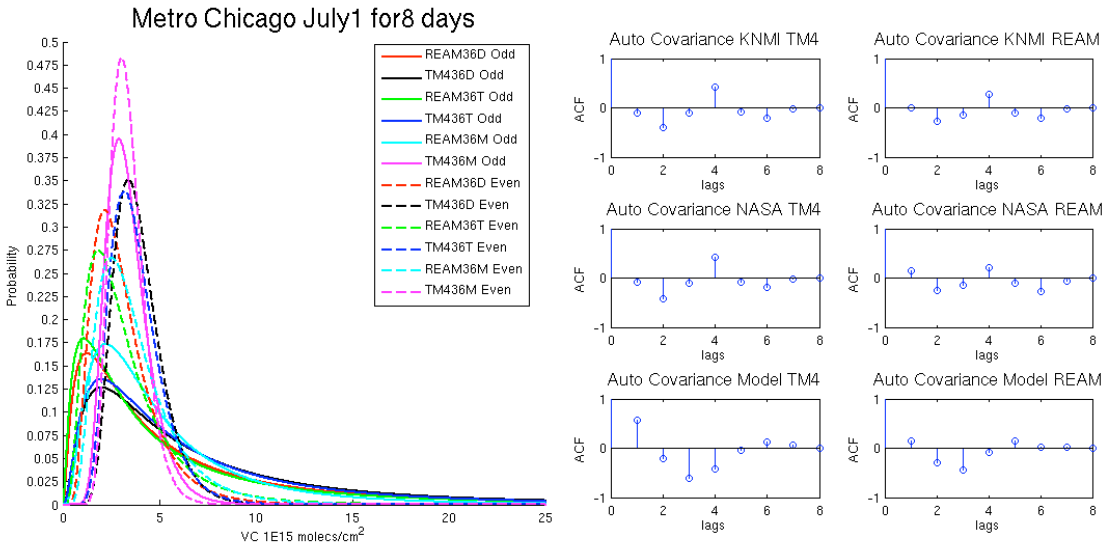


Figure 3.6.4: Distribution plot for Chicago starting July 1st going for 8 days

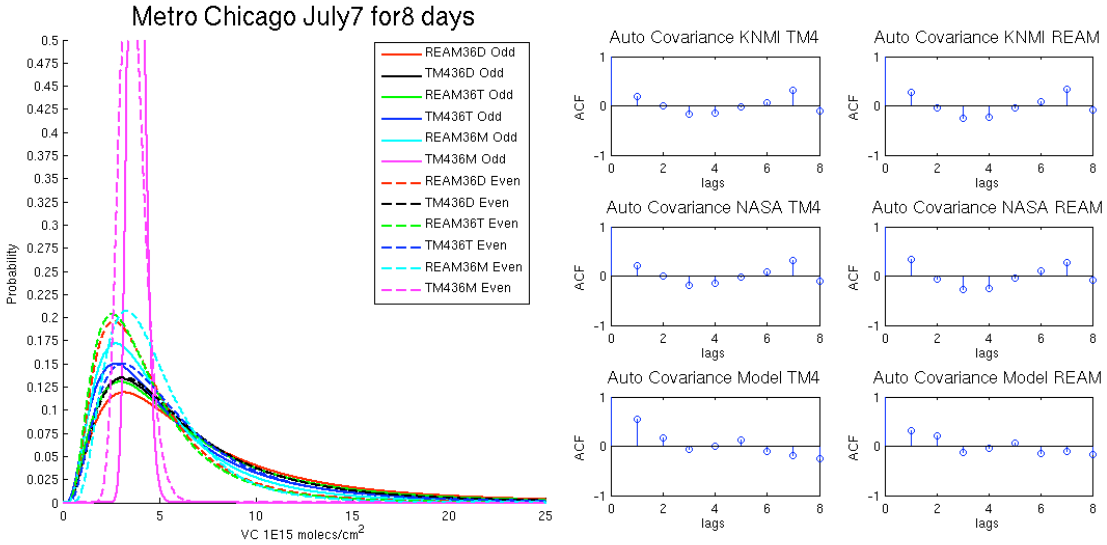


Figure 3.6.5: Distribution plot for Chicago starting July 7th going for 8 days

Chicago was chosen to demonstrate more clearly the effects of port and starboard biases. Chicago NOx chemistry is much better represented by regional models than that of Atlanta, furthermore making Chicago a more suitable testing ground. Referring to Figure 3.6.2 Chicago is port of the satellite OMI on odd days. It is observed that all the retrievals on odd days indicate higher NO_2 tropospheric vertical column means as longer log-normal tails. Even days on the other hand are cleaner. Referring to Table A.2, only the TM4 mean on odds days compared with the KNMI TM4 and NASA TM4 retrieval is dissatisfactory. On the other hand, REAM model means are always satisfactory compared with retrieval means. The only model that passed the f-test is REAM on even days compared with the NASA REAM retrieval. Despite TM4 retrieval coupling, REAM auto-covariance agrees better with the auto-covariance of all four retrievals.

On the other hand, when the same analysis was performed on the same city for a period 7 days later, REAM agreed much better with the retrieval. For example, on July 7^{th}, Chicago starts just starboard of nadir. Being starboard and near nadir reduces the effect of the relative azimuth bias. Referring to Table A.2, both TM4 and REAM manage the capture the mean NO_2 tropospheric vertical column well. However, with this near nadir setup, REAM now out-performs TM4 by two orders of magnitude on the f-test, although it only comes close to passing the test.

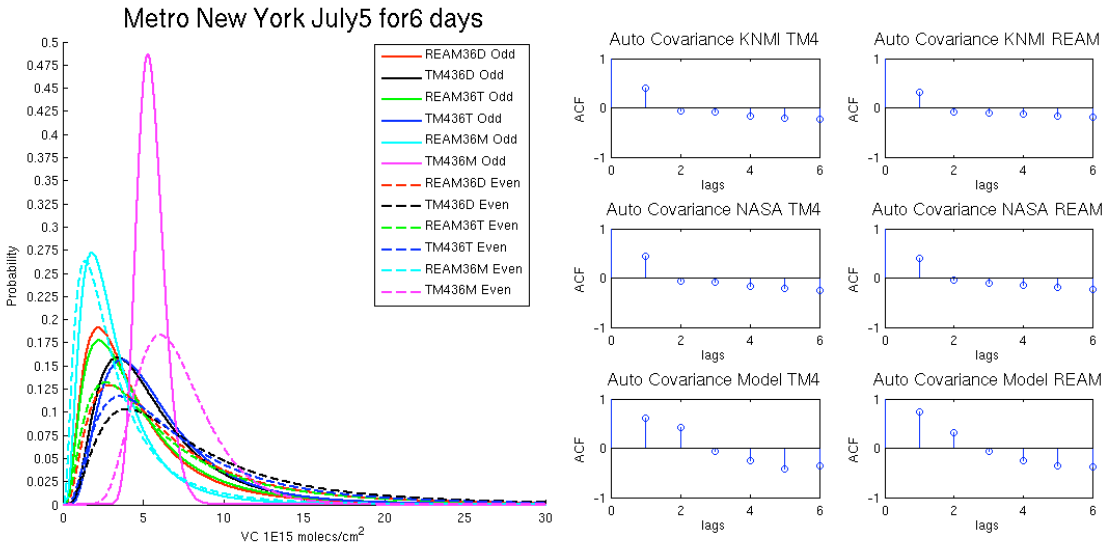


Figure 3.6.6: Distribution plot for New York City starting July 5th going for 6 days

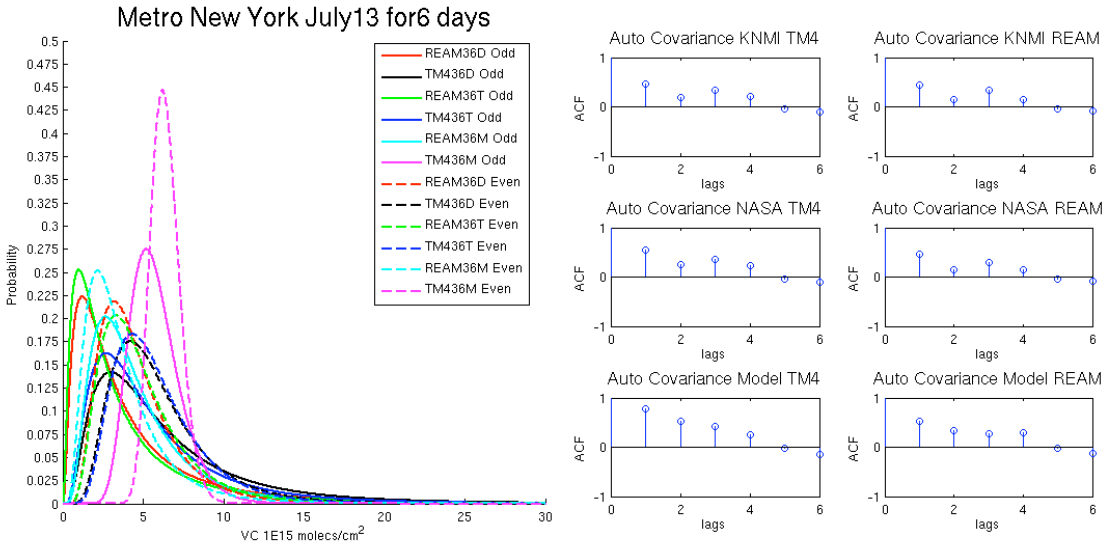


Figure 3.6.7: Distribution plot for New York City starting July 13th going for 6 days

New York City was selected to show the adverse effect of coupling the chemical transport model to the retrieval. On July 5th New York City is about as far off nadir as possible just starting on the starboard side of the swath. Thus, it would be expected that odd days are less polluted with NO_2 than even days. This is the case, though for the interval in question New York City is approaching nadir from the starboard side; thus experiencing minimal effects of the relative azimuth angle bias. It is important to note, however, that TM4 had developed a strong bipolar moment between the even and odd days. This should be compared to REAM which was relatively stable from day to day being that it is independent of the retrieval. REAM is nevertheless a little bit low, as indicated in Figure 3.6.5 and the t-tests in Table A.4. As a consequence of this bipolar moment, TM4 for once does better than REAM on the f-test on even days whereas on odds REAM assumes superiority again. This TM4 bipolar moment is reversed on the 8 days leaving nadir where New York City and is found on the port side of the satellite on odd days. This is the converse situation as before and likewise the TM4 even day-odd day polarity has switched.

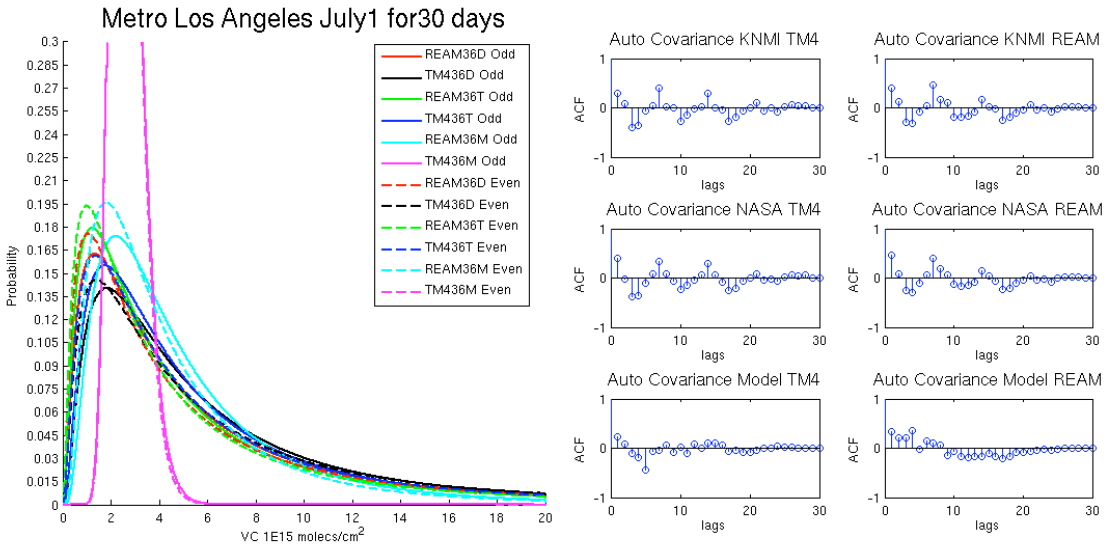


Figure 3.6.8: Distribution plot for Los Angeles starting July 1st going for 30 days

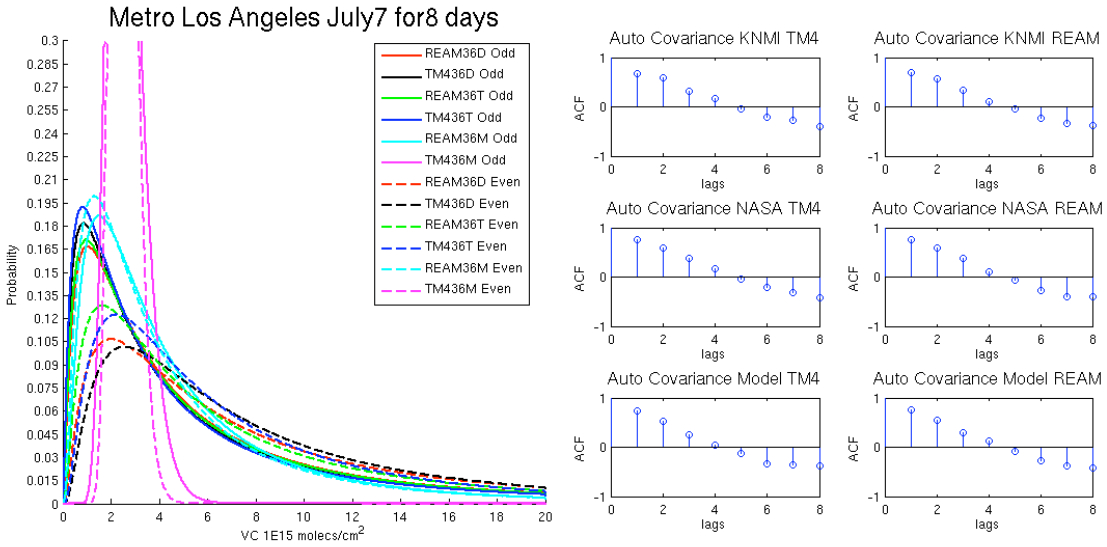


Figure 3.6.9: Distribution plot for Los Angeles starting July 7th going for 8 days

Los Angeles was selected because of all these cities, it experiences minimal cloudy days in July. Figure 3.5.8 shows Los Angeles from the 1st of July to the 30th of July. These satellite effects, when averaged out over a month, disappear. Even for the monthly average, REAM vastly out-performs TM4 and NASA retrievals agree better with the models. Only the NASA REAM retrieval mean NO_2 tropospheric column concentration agrees within the 95% confidence interval with the REAM model. All other tests fail. Nevertheless, despite the long duration of this analysis, these retrieval periodic effects retain their footprint in the auto covariance. All four of the retrievals show the eight-day period that was explored in depth with the previous cities. When observing a whole month together, the effect of TM4 coupling with the retrieval is made evident. TM4 coupling lags behind the retrieval by as much as one day at times. REAM, on the other hand, does not show any such dependency. In all cases, after one month, a point on the series is no longer correlated with point further out than one month.

In narrowing down the interval to eight days, this created a scenario in which on the odd days Los Angeles is just leaving one swath and starting the starboard side of the other swath and on the even days Los Angeles is leaving nadir and starting port side of the swath. As expected, KNMI retrieval produces a highly polluted distribution. However, this distribution is likely too high since the TM4 model does not even come close to agreeing with its own retrieval. As was seen in the previous case, Figure 3.6.8, REAM does better all around, and the NASA script agrees better with the retrieval (Table A.7).

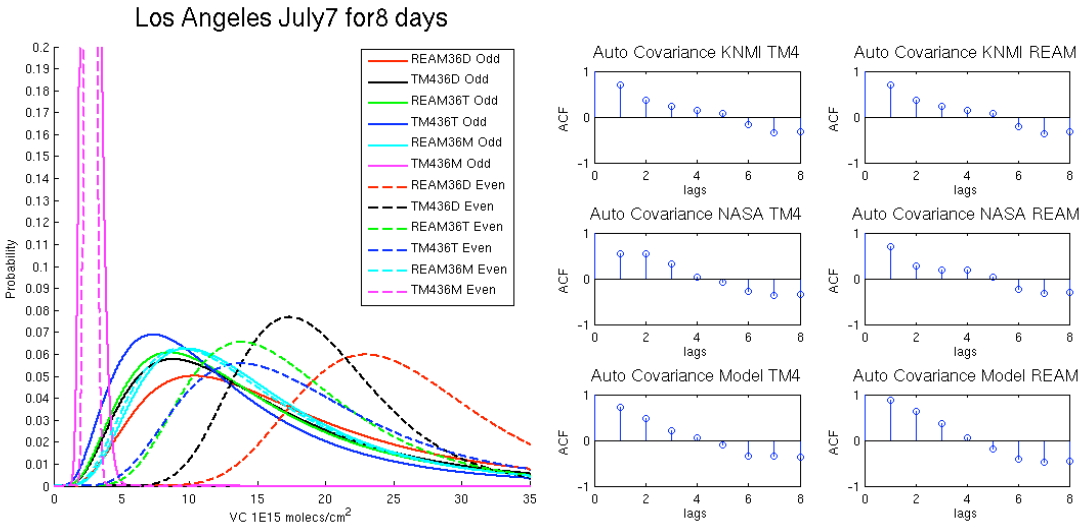


Figure 3.6.10: Distribution plot for inner Los Angeles using REAM36 domain.

Figure 3.6.10 shows the spatial variability of NO_2 compared to the two previous Los Angeles figures, Figures 3.6.8 and 3.6.9. Here the REAM36DAK and the TM436DAK retrievals for even days indicated by the dashed red and black lines, respectfully, are excessive. The DAK radiation model is sensitive to the REAM36 a priori model and its retrieval mean is significantly different than the other retrievals (Table A.8).

CHAPTER 4

DISCUSSION

The original hypothesis is that there would be no significant difference between the retrievals using different a priori models or different radiative models. The results show, however, that not only are there significant differences in few of the means and most of the model a priori distributions calculated here, but there are significant problems as well.

Over New York City, for example, TM4 observes higher NO_2 levels on even days and low NO_2 levels on odd days. The model and retrieval are coupled such that this generates a positive feedback mechanism where TM4 will predict higher NO_2 on even days and lower NO_2 on odd days (Figure 3.6.5) until finally odd and even days are viewed on opposite sides of the satellite and the reverse positive feed mechanism occurs, Figure 3.6.6. Under such circumstances, TM4 becomes less stable and reliable.

The most highly unexpected result is the low vertical columns of the retrievals using REAM. The reason this was at first unexpected is because the REAM model has more NO_2 in the vertical column than the TM4 model. Nevertheless, the result of the retrieval only depends only on the distribution of NO_2 in the vertical column. REAM has more NO_2 aloft to due lightning production. This NO_2 aloft increases the total tropospheric slant column concentration which is proportional to the tropospheric air mass factor, and thus reduces tropospheric vertical column. This implies either that

REAM has too much NO_2 aloft or that the radiation model is too sensitive to NO_2 aloft and in either case the REAM model does not agree well with observations in rural areas.

Another entirely unexpected result is the dependency of the air mass factor on the relative azimuth angle. In other words, KNMI is more sensitive to locations to the west of the satellite path, port, and less sensitive to locations to the east of this overpass, starboard. For a city on the ground, being west or east switches every day. This is related to the radiation model used by KNMI and to a lesser extent by NASA. It is possible that the satellite's sensor asymmetry has changed since being first launched and the radiation tables simply need to be reoptimized for different viewing geometries. Nevertheless, for information regarding NO_2 vertical columns, this day-to-day effect can be quite dramatic as was seen in the case for New York City. This should call into question the "Near Real Time" application promoted KNMI since rough 3 days need to be averaged to reduce the effect of the problem.

However, most of the significant differences between the means and distributions of NO_2 in the model and from the retrievals were expected. These differences are mainly due to the on horizontal resolution of REAM and the known sensitivity of KNMI to large, near-surface NO_2 sources. These differences were to be expected as result of previous studies and known advantages of using higher resolution chemical transport models.

As was seen in Figure 1.4.3, in the INTEX-B flight shows that NASA averages slightly less than as observed by the flight and KNMI retrieval average over with greater NO2 than the NO_2 columns observed in the flight. Although this study looks at a

different time and with no direct ground or air validation, the same results are found to persist here: KNMI averages greater NO_2 tropospheric vertical columns than NASA. If the assumption that NASA, though under predicting, is closer to the validation than KNMI persists as well, then it would be possible to conclude that REAM modeled profiles with the NASA modeled radiation does best to predict NO_2 in urban regions.

There is, however, a severe implication on the role of lightning in this retrieval. REAM is engineered to match lightning observations and has more lightning NOx production then TM4. It is observed directly how much larger NO_2 aloft in REAM that can be compared to levels in TM4. In mixing ratio terms, this represents more than 200% more NO_2 in REAM. High slant columns, as determined by the radiation model, give high air mass factors and thus a lower retrieval vertical column. This NO_2 aloft in REAM severely redistributes the NO_2 altitude-dependent air mass factor that subsequently underestimates the NO_2 in rural areas relative to TM4. REAM's urban pollution increase off-setd this effect. Nevertheless, comparing the REAM model to the REAM retrieval, the opposite is found: REAM over predicts NO_2 in rural areas and under predicts the cities. An even higher resolution REAM model that constricts NO_2 to the cities may fix this problem, agree more closely with the ground observations and adjust our thinking of NO_2 emissions.

CHAPTER 5

CONCLUSIONS

Tests on the radiation models and chemical transport models demonstrate potential problems which confound the retrieval. One issue involves the radiation model include a port and starboard satellite bias which make KNMI more sensitive to near surface NO_2 on the port side. Issues as result of the chemical transport model include problems regarding the distribution of NO_2 aloft in the troposphere, inaccurate emissions inventories, and sensitivity regarding horizontal resolution.

In all four North American cities tested, Atlanta Los Angeles, New York City, and Chicago, both REAM and TM4 failed the f-test for getting the variability of NO_2 columns over that city correct. There were a couple of exceptions where REAM and the retrieval agreed on variability. Aside from the city of Atlanta, REAM and TM4 did fairly well at getting the mean concentration of NO_2 over cities accurate. The exception was over Los Angeles where TM4 failed the T-test but REAM did not. Overall, REAM, a regional chemical transport model, did much better at matching all retrieval outputs. In general, however, the retrieval using TOMRAD outperformed the DAK retrieval in agreeing with the REAM model and in producing balanced day-to-day results. This does not imply that the NASA retrieval correlates better with ground measurements than KNMI, but only that it agrees better with the models, TM4 and REAM.

There are several avenues for further study regarding the retrieval and its implications on our understanding of air quality. These suggestions could lead to a very successful, perpetuating, Ph.D. thesis and a continued sustained research group within Earth and Atmospheric Science at Georgia Tech.

The first logical step would be to reinvestigate the radiative model underlying the retrieval. There is beyond a doubt a statistical difference between air mass factors computed on the east and west sides of the satellite. Hence, coupling the chemical transport model to the retrieval makes the chemical transport model more unstable. Such an investigation should aim to understand how these differences arise and correct for these differences such that there is no longer a robust systematic error attributed to the relative azimuth angle. This step would affect all future research and application of these two OMI products evaluated here, KNMI and NASA, and therefore should receive first priority. This would require further collaboration with KNMI and NASA to acquire a deeper understanding of their radiative transfer models.

It would also be useful to investigate adding a water vapor dimension to the radiation look up table. Such a variable could account for water vapor in the slant column, hydroscopic aerosols, optically thin clouds, and perhaps even improve upon the cloud fraction restriction. Since it would be calculated off-line, it would not cost the retrieval in terms of performance.

The constant profile test needs to be rerun using REAM a priori profiles at certain extremes. Such extremes would account for high near-surface NO_2 at different pressures and albedos (held previously constant), and one high lightning NO_2 aloft profile. The TM4 profiles used in this thesis just barely were within the published error ranges.

In addition, it would be useful to have NO_2 soundings for a period of time over a particular region of interest in order to validate chemical transport models and their retrievals. Currently, field data typically includes, one or two airborne studies and a couple of scattered column measurements. Regarding the problems of this retrieval presented here, it is impossible to say if the model is correctly assessing NO_2 via chemical transport model assumptions or if the retrieval is disproportionally weighting NO_2 when computing air mass factors. Light aircraft could be solicited from around the country to take sample profiles for a period of one month under visual flight rules at the time of satellite overpass. Six sites operated over the span of a month could be quite informative regarding model and retrieval validation.

It would be useful to take the minimum, maximum, and average case retrievals, and perform inversions to improve current emissions inventories. Furthermore, although REAM did an excellent job in capturing the variability of NO_2 over cites of interests compared to TM4, it still more often than not produces distributions statistically different than that of the retrieval. This is observed even in the monthly average. The cause for this is the incorrect distribution of NO_2 in the vertical column. These models require

further optimization to better understand air quality and physical and chemical processes.

In connection with the above idea about improving the models, a good model and a good retrieval could be used to back-out cloud albedo. The current algorithm assumes that all clouds at all pressures have an albedo of 0.8. This seems to be an awfully basic assumption to make, which could have a profound effect on determining the amount of NO_2 over a partly clouded or completely clouded cell.

Another avenue for potential study would be to develop a recursive retrieval algorithm. This would work by taking the a priori guess, calculating the retrieval and then adjusting the a priori guess intelligently, taking the retrieval again, and repeating the process until an acceptable level of convergence is attained. Such retrieval would rely on accurate radiation models and a stratospheric model to first estimate the tropospheric vertical column, as should be done now. Referring to section 1.5 for ideas on NO_2 source separation, one would then access the spatial and temporal observed effects of NO_2 in the tropospheric column. If the feature is wide spread with a lifespan of several days, it is probably NO_2 as result of lightning in the mid to upper troposphere. If the feature is concentrated and variable, then it is probably an anthropogenic source of NO_2 and near the surface. In making these two assumptions, the retrieval could actively adjust the vertical column distribution of the a priori guess to best suit its *own* observations.

The more lightning the model uses, the smaller retrieval result. This would confound soil emissions, especially in rural regions. How does the inversion know what is lightning

and what is soil? This is why I suggested a recursive retrieval/inversion script. If the script understands how NO_2 evolves, it can distinguish the seasonal NO_2 soil cycle from the synoptic meteorology input and adjust until the model and retrieval are within error tolerance.

APPENDIX

Table A.1: Statistics for Atlanta starting on July 1st going for 30 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	5.0589	0.6262	8.3429	1.285	1.5263e-10	0.00026648	5.1576e-08	0.050655
model-nasaTM4	4.7888	0.5509	8.1488	1.2362	1.0587e-08	0.0070286	3.6202e-06	0.49543
model-nasaREAM:	4.4805	0.4649	7.334	1.0311	1.1457e-06	0.13692	0.00012935	0.93502
model-knmiREAM:	6.7965	1.111	10.3441	1.7887	1.9095e-08	0.0012128	9.163e-07	0.018891

Retrieval vs Retrieval	f-odd	f-even	t-odd	t-even
REAMnasa/knmi:	0.3272	0.2687	0.050826	0.028642
TM4nasa/knmi	-0.0427	-0.0677	0.3329	0.25099
nasaTM4/REAM:	-0.0333	0.0081	0.22573	0.51896
knmiTM4/REAM:	0.194	0.1235	0.99681	0.44389
modelTM4/REAM:	2.4961	2.8846	8.2253e-05	8.6646e-08
modelTM4/REAM:	1.9569	3.5881	0.90413	0.0031513

Even vs Odd	f-test	t-test
knmiTM4:	0.0873	0.058668
knmiREAM:	0.0227	0.43758
nasaTM4:	0.116	0.037467
nasaREAM:	0.0706	0.12913
modelTM4:	0.2159	0.0031513
modelREAM:	0.0884	0.90413

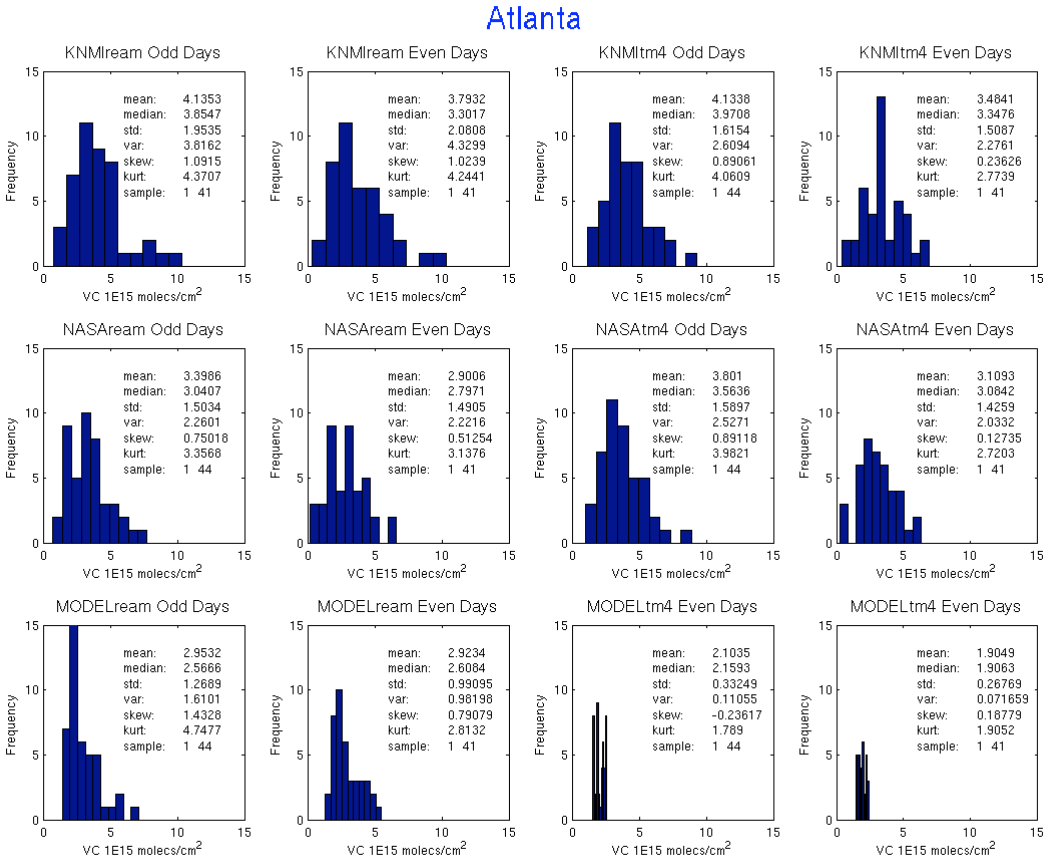


Figure A.1: Histogram for Atlanta starting on July 1st going for 30 days

Table A.2: Statistics for Atlanta starting on July 1st for 30 days no lightning

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	5.0589	0.783	8.3429	1.4907	1.5263e-10	1.3624e-06	5.1576e-08	0.00092124
model-nasaTM4	4.7888	0.7007	8.1488	1.4376	1.0587e-08	7.7106e-05	3.6202e-06	0.03014
model-nasaREAM:	6.0231	1.0765	10.438	2.0637	6.0439e-10	1.811e-06	1.2359e-07	0.0008145
model-knmiREAM:	16.3789	4.2294	18.3526	4.2282	4.8546e-11	1.7755e-09	9.3977e-11	2.1364e-08

Retrieval vs Retrieval	f-odd	f-even	t-odd	t-even
REAMnasa/knmi:	1.3675	0.5979	0.00076814	0.00056368
TM4nasa/knmi	-0.0427	-0.0677	0.3329	0.25099
nasaTM4/REAM:	0.1213	0.1591	0.24942	0.18545
knmiTM4/REAM:	1.7527	0.9726	0.00036811	0.00023475
modelTM4/REAM:	2.1959	2.5678	0.018023	0.00036312
modelTM4/REAM:	1.714	3.1964	0.85293	0.0031513

Even vs Odd	f-test	t-test
knmiTM4:	0.0873	0.058668
knmiREAM:	0.0835	0.16144
nasaTM4:	0.116	0.037467
nasaREAM:	0.1537	0.082474
modelTM4:	0.2159	0.0031513
modelREAM:	0.0832	0.85293

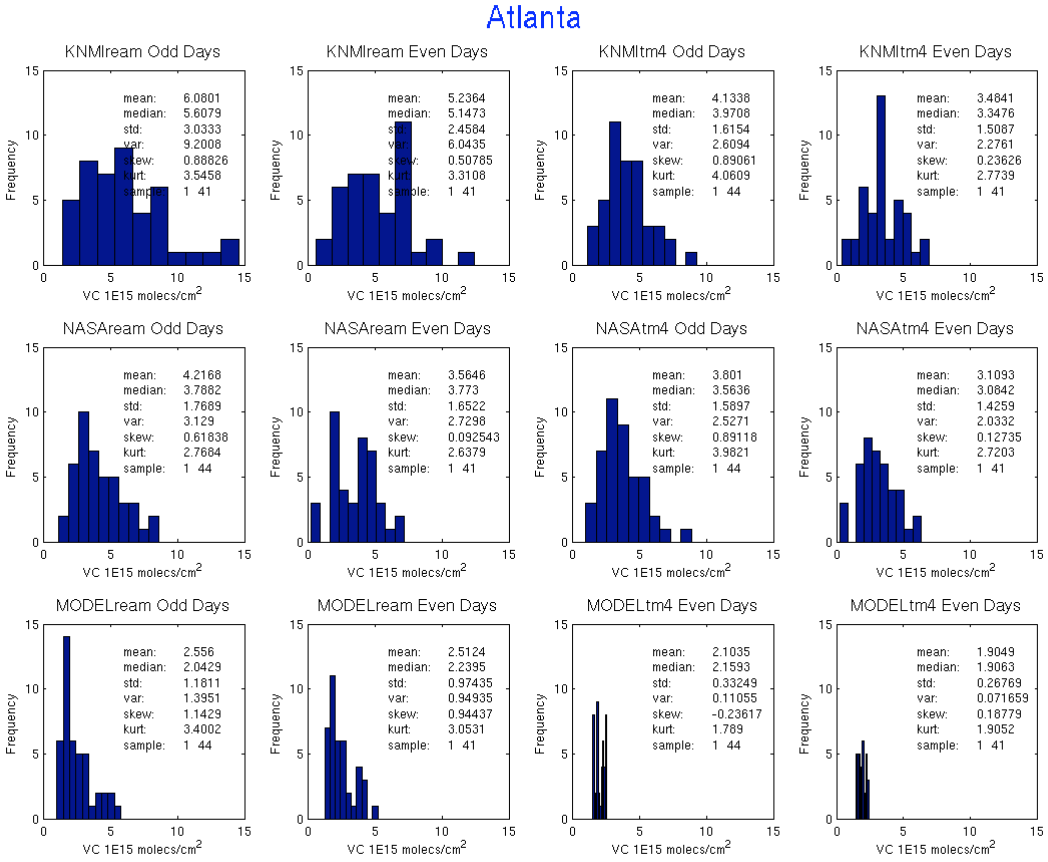


Figure A.2: Histogram for Atlanta starting on July 1st for 30 days no lightning

Table A.3: Statistics for Atlanta starting on July 5th going for 8 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	5.7588	1.3609	3.6237	0.3598	0.00053163	0.00048392	4.7804e-05	2.8513e-05
model-nasaTM4	5.1757	1.1511	3.4636	0.3101	0.0021696	0.002062	1.2816e-05	9.4898e-06
model-nasaREAM:	4.2758	0.8275	4.9907	0.7841	0.049866	0.040905	0.012147	0.0046037
model-knmiREAM:	6.5313	1.6387	5.9186	1.0722	0.0010681	0.00091228	0.0017425	0.00063975

	f-test		t-test	
Retrieval vs Retrieval	odd	even	odd	even
REAMnasa/knmi:	0.3091	0.0424	0.11795	0.41774
TM4nasa/knmi	-0.0169	-0.0747	0.5737	0.86022
nasaTM4/REAM:	0.0574	0.2242	0.26563	0.29682
knmiTM4/REAM:	0.0028	0.3751	0.98851	0.94156
modelTM4/REAM:	1.6704	2.1115	0.53869	0.23226
modelTM4/REAM:*	2.0512	1.7235	0.55662	0.87895

Even vs Odd	f-test	t-test
knmiTM4:	0.3691	0.99134
knmiREAM:	-0.0018	0.93023
nasaTM4:	0.2909	0.41713
nasaREAM:	0.0018	0.44021
modelTM4:	-0.0906	0.87895
modelREAM:	0.0273	0.55662

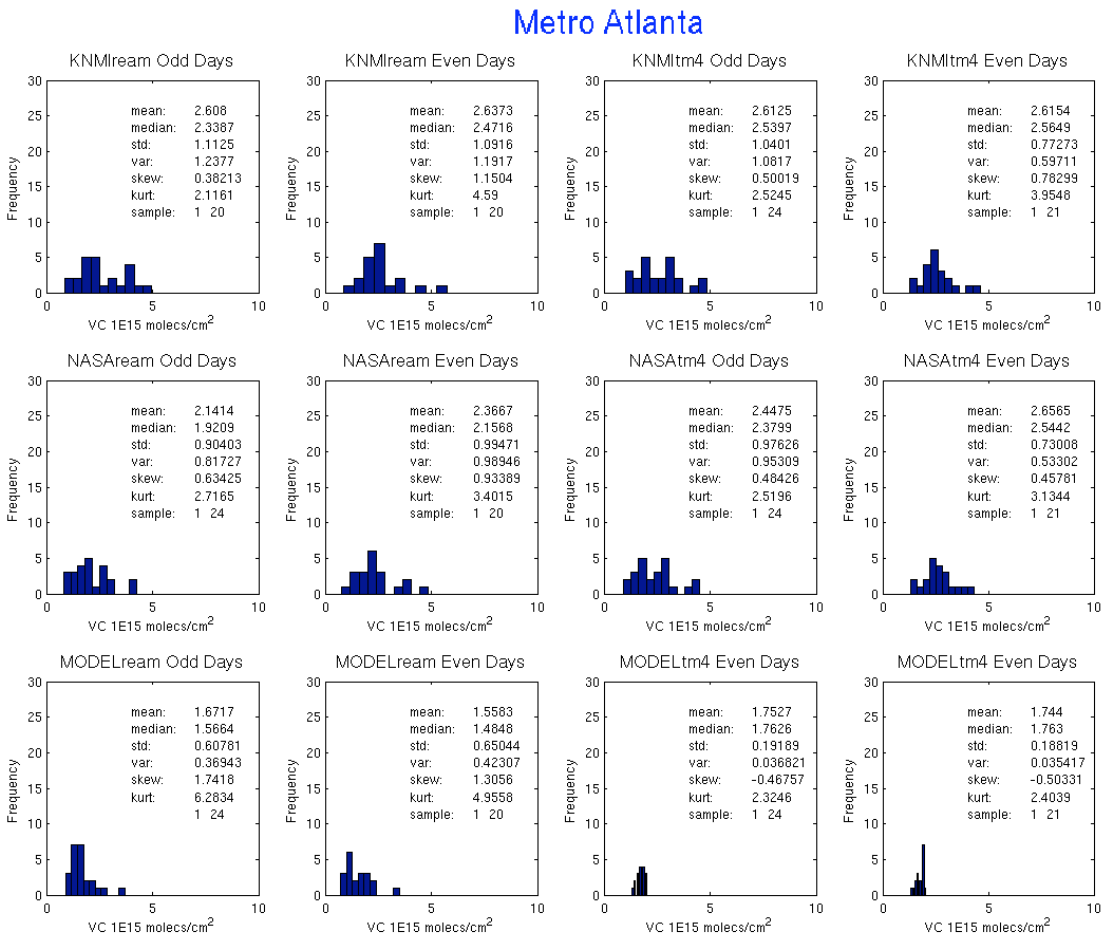


Figure A.3: Histogram for Atlanta starting on July 5th going for 8 days

Table A.4: Statistics for Chicago starting on July 1st going for 8 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	5.6406	0.8568	0.392	0.4176	0.035236	0.23662	0.1809	0.68012
model-nasaTM4	4.7282	0.595	0.4697	0.3425	0.043046	0.33697	0.31651	0.85081
model-nasaREAM:	3.0103	0.1021	1.1146	-0.0617	0.11209	0.65139	0.9039	0.57598
model-knmiREAM:	3.5762	0.2645	0.6611	0.1863	0.085827	0.50182	0.65557	0.39858

	f-test		t-test	
Retrieval vs Retrieval	odd	even	odd	even
REAMnasa/knmi:	0.0707	0.1942	0.82983	0.82552
TM4nasa/knmi	0.0894	-0.0148	0.78304	0.79616
nasaTM4/REAM:	0.3532	0.3514	0.64892	0.42721
knmiTM4/REAM:	0.3765	0.1163	0.62927	0.16054
modelTM4/REAM:	2.4173	1.1009	0.10205	0.54096
modelTM4/REAM:*	0.6329	3.3764	0.19593	0.9603

Even vs Odd	f-test	t-test
knmiTM4:	4.7913	0.06266
knmiREAM:	2.3569	0.069822
nasaTM4:	3.7387	0.071374
nasaREAM:	1.3147	0.11138
modelTM4:	0.2071	0.9603
modelREAM:	0.9811	0.19593

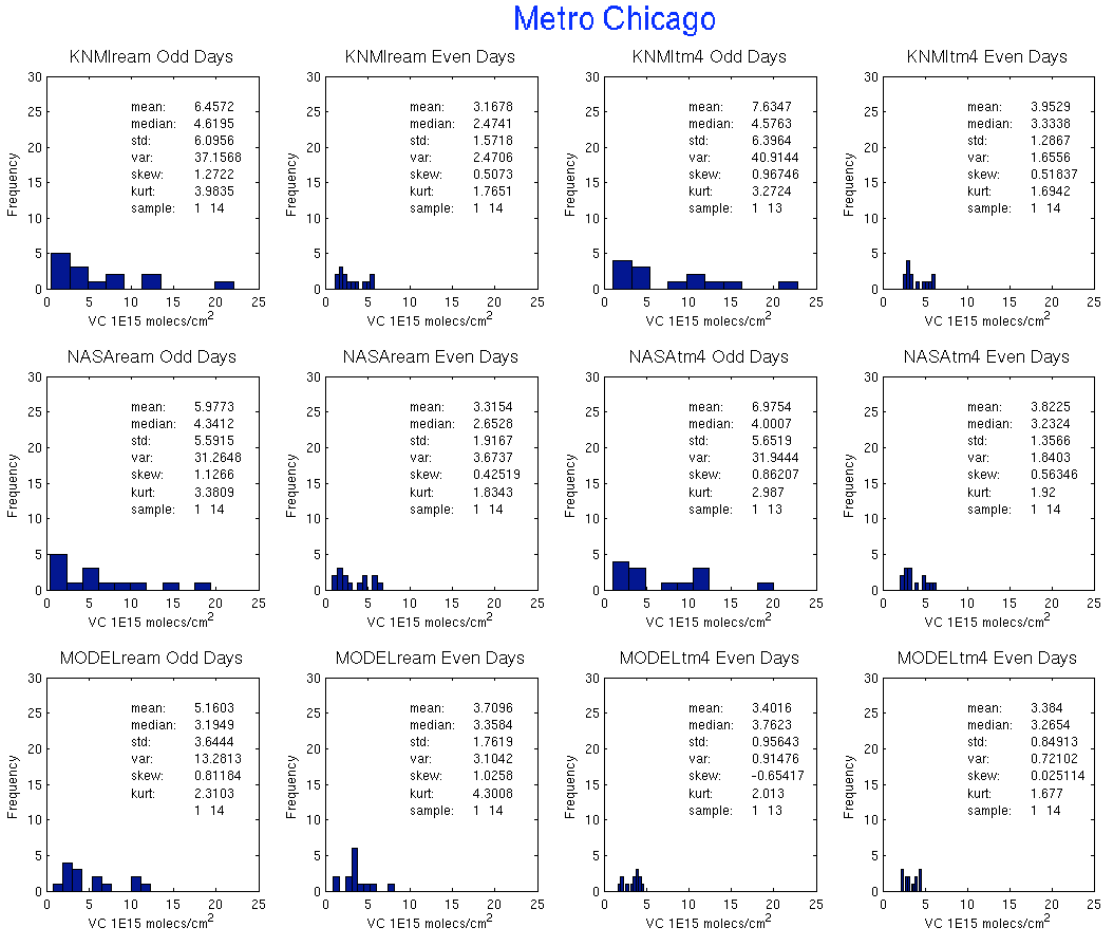


Figure A.4: Histogram for Chicago starting on July 1st going for 8 days

Table A.5: Statistics for Chicago starting on July 7th going for 8 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	14.69	0.499	7.6283	1.0677	0.10652	0.41673	0.12999	0.34133
model-nasaTM4	11.9165	0.2232	6.1078	0.6943	0.15743	0.61978	0.14762	0.44615
model-nasaREAM:	15.6337	0.5929	3.2995	0.0046	0.085123	0.37927	0.33146	0.82752
model-knmiREAM:	18.9217	0.9198	3.558	0.0681	0.069158	0.26777	0.2685	0.95229

	f-test		t-test	
Retrieval vs Retrieval	odd	even	odd	even
REAMnasa/knmi:	0.1288	-0.0088	0.77238	0.8873
TM4nasa/knmi	0.1456	0.1439	0.74956	0.79848
nasaTM4/REAM:	0.2183	0.575	0.71882	0.38398
knmiTM4/REAM:	0.2005	0.8118	0.75605	0.33869
modelTM4/REAM:	8.9879	3.0038	0.18206	0.15608
modelTM4/REAM:*	6.0533	4.6817	0.71712	0.3185

Even vs Odd	f-test	t-test
knmiTM4:	-0.0319	0.9824
knmiREAM:	1.4027	0.1787
nasaTM4:	-0.0334	0.9879
nasaREAM:	1.1185	0.21042
modelTM4:	0.6809	0.3185
modelREAM:	0.3441	0.71712

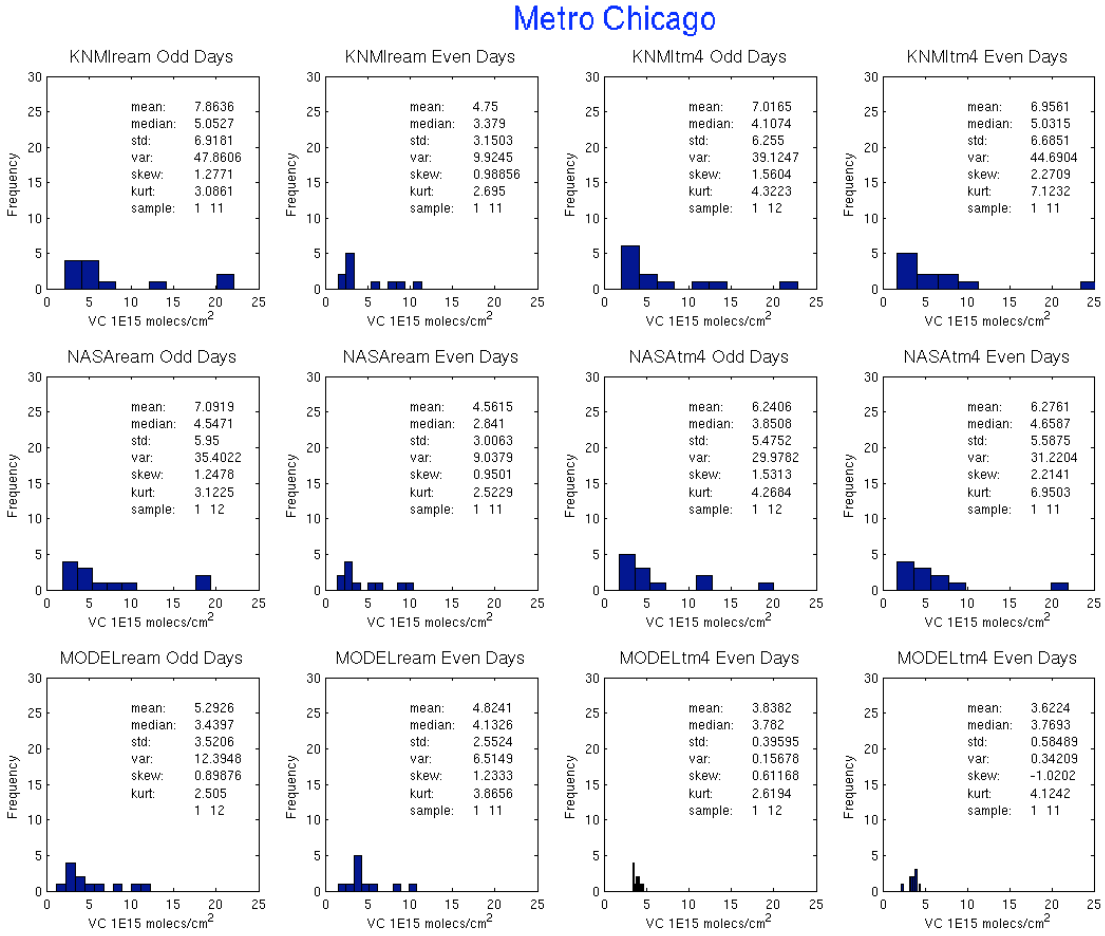


Figure A.5: Histogram for Chicago starting on July 7th going for 8 days

Table A.6: Statistics for New York City starting on July 5th going for 6 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	3.4015	0.9592	1.8481	2.8961	0.76692	0.063637	0.48266	0.026587
model-nasaTM4	3.4566	0.9842	1.2619	2.0987	0.64539	0.048387	0.78906	0.040258
model-nasaREAM:	3.0056	0.7799	0.9163	1.6287	0.7471	0.20664	0.84716	0.077362
model-knmiREAM:	2.5592	0.5778	0.9561	1.6828	0.49625	0.27265	0.92098	0.069131

	f-test		t-test	
Retrieval vs Retrieval	odd	even	odd	even
REAMnasa/knmi:	0.0614	-0.0418	0.81865	0.94277
TM4nasa/knmi	-0.0496	0.1904	0.90245	0.70426
nasaTM4/REAM:	0.049	0.1128	0.56893	0.71825
knmiTM4/REAM:	0.1707	0.3802	0.47324	0.5141
modelTM4/REAM:	1.1465	0.2982	0.0076738	0.0017388
modelTM4/REAM:*	1.3887	0.4474	0.95121	0.058726

Even vs Odd	f-test	t-test
knmiTM4:	1.1112	0.21935
knmiREAM:	0.7956	0.26687
nasaTM4:	0.6522	0.42348
nasaREAM:	0.5592	0.40186
modelTM4:	2.2712	0.058726
modelREAM:	0.0478	0.95121

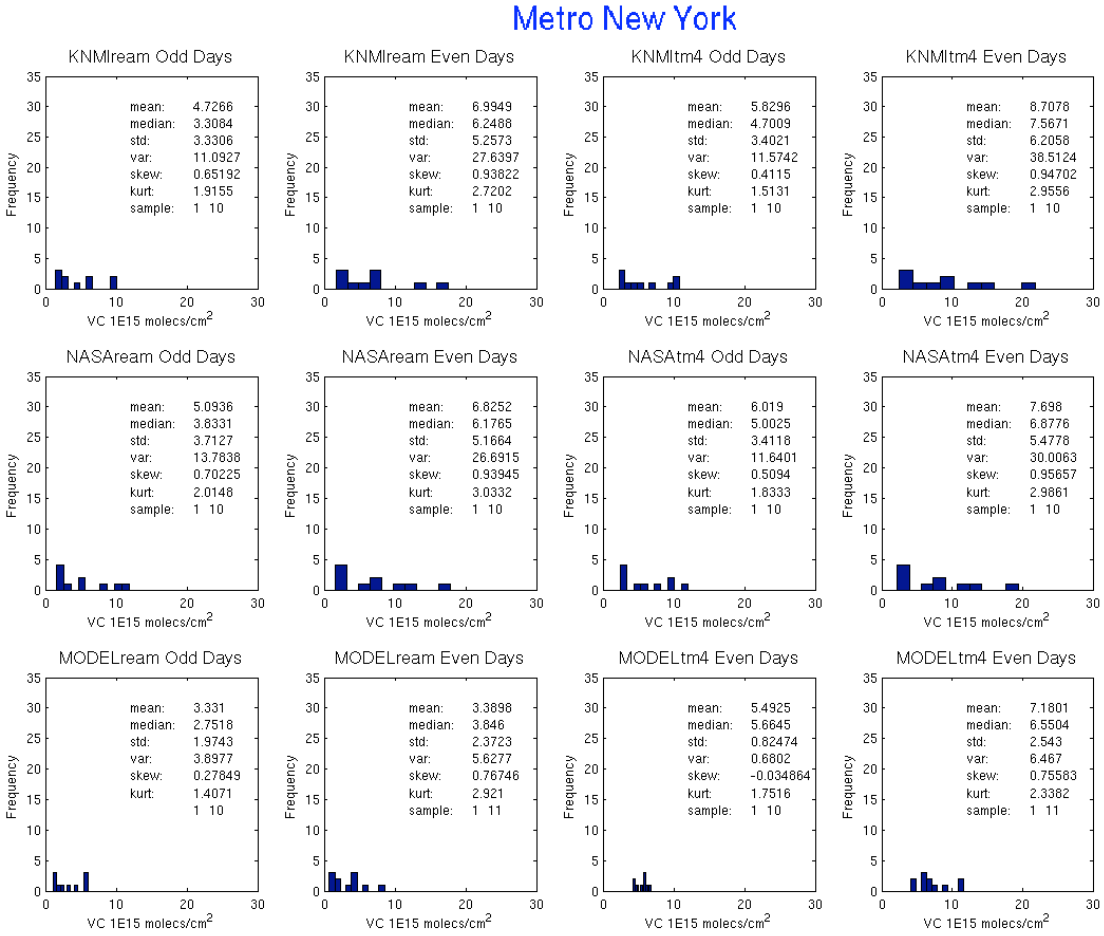


Figure A.6: Histogram for New York City starting on July 5th going for 6 days

Table A.7: Statistics for New York City starting on July 13th going for 6 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	1.7914	0.6875	2.2929	0.4789	0.60684	0.30169	0.54466	0.054412
model-nasaTM4	1.2443	0.3521	2.0701	0.3766	0.91455	0.47846	0.48187	0.052811
model-nasaREAM:	0.3967	0.0563	1.7163	0.2141	0.18932	0.69816	0.053086	0.29106
model-knmiREAM:	0.6144	-0.0341	1.4536	0.0934	0.41226	0.97406	0.0092373	0.40618

	f-test		t-test	
Retrieval vs Retrieval	odd	even	odd	even
REAMnasa/knmi:	0.0874	0.0425	0.78331	0.74217
TM4nasa/knmi	0.1753	0.0092	0.74638	0.96277
nasaTM4/REAM:	0.5193	0.0654	0.36542	0.32726
knmiTM4/REAM:	0.6402	0.2717	0.34941	0.17147
modelTM4/REAM:	0.5691	1.1154	0.20727	0.0040825
modelTM4/REAM:*	0.1602	1.8438	0.44845	0.22233

Even vs Odd	f-test	t-test
knmiTM4:	0.4533	0.66751
knmiREAM:	0.1253	0.9924
nasaTM4:	0.25	0.95361
nasaREAM:	0.0072	0.55643
modelTM4:	0.7197	0.22233
modelREAM:	0.2727	0.44845

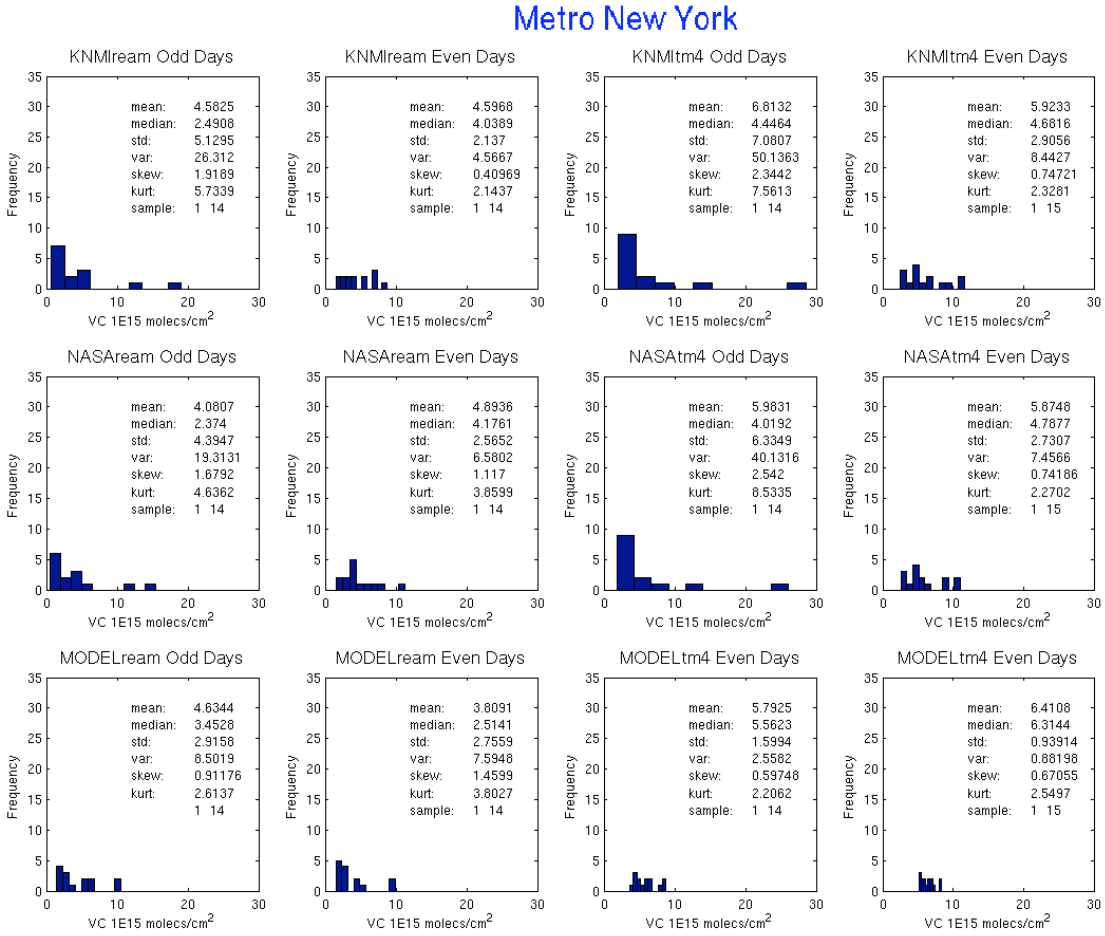


Figure A.7: Histogram for New York City starting on July 13th going for 6 days

Table A.8: Statistics for Los Angeles starting on July 1st going for 30 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	10.643	0.5854	11.2611	0.9075	2.2156e-09	0.01796	1.0819e-07	0.002656
model-nasaTM4	8.4779	0.2784	8.9533	0.5364	1.0034e-08	0.13818	4.2269e-07	0.02243
model-nasaREAM:	6.866	0.0498	6.5935	0.157	1.0585e-06	0.29852	1.0654e-05	0.15082
model-knmiREAM:	8.368	0.2628	8.1036	0.3998	1.6079e-07	0.08079	1.9593e-06	0.033619

	f-test		t-test	
Retrieval vs Retrieval	odd	even	odd	even
REAMnasa/knmi:	0.1128	0.1205	0.50528	0.48768
TM4nasa/knmi	0.1503	0.1537	0.34955	0.38193
nasaTM4/REAM:	0.1266	0.2313	0.76905	0.44101
knmiTM4/REAM:	0.1645	0.2676	0.64947	0.36812
modelTM4/REAM:	5.9761	5.1424	1.1321e-08	3.6316e-06
modelTM4/REAM:*	4.8252	6.3552	0.31876	0.78425

Even vs Odd	f-test	t-test
knmiTM4:	-0.0754	0.69936
knmiREAM:	0.0078	0.8751
nasaTM4:	-0.0725	0.76461
nasaREAM:	0.0148	0.80448
modelTM4:	-0.0226	0.78425
modelREAM:	0.1187	0.31876

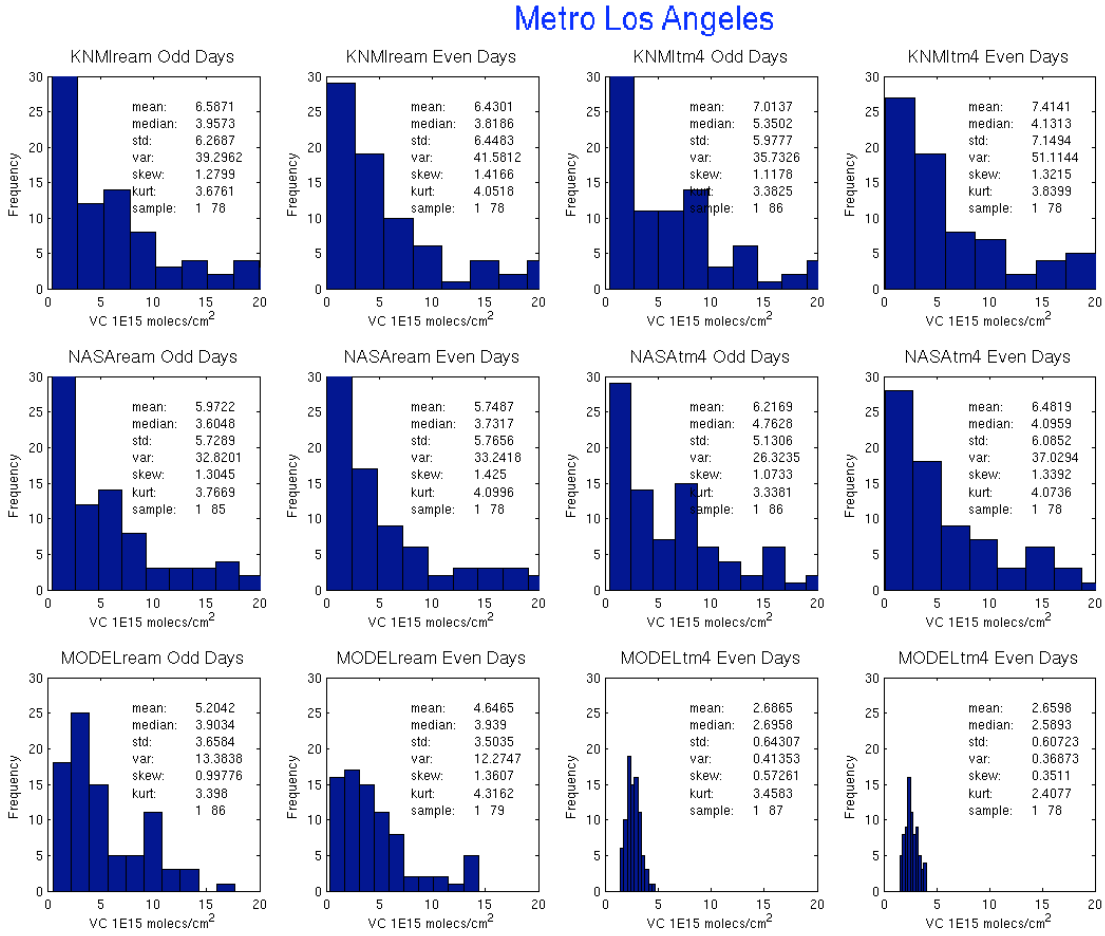


Figure A.8: Histogram for Los Angeles starting on July 1st going for 30 days

Table A.9: Statistics for Los Angeles starting on July 7th going for 8 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	6.7176	0.1941	30.1423	2.5578	0.0064123	0.2911	0.00054686	0.019256
model-nasaTM4	6.009	0.0786	19.5661	1.3266	0.0084232	0.40016	0.00092427	0.070539
model-nasaREAM:	7.3723	0.3008	19.0677	1.2686	0.006801	0.24525	0.0009723	0.066484
model-knmiREAM:	7.6562	0.3471	28.458	2.3617	0.0057302	0.23428	0.00045217	0.01858

	f-test		t-test	
Retrieval vs Retrieval	odd	even	odd	even
REAMnasa/knmi:	-0.0428	0.3898	0.99391	0.48512
TM4nasa/knmi	0.0237	0.436	0.82484	0.43328
nasaTM4/REAM:	0.1161	-0.0516	0.72001	0.95789
knmiTM4/REAM:	0.0441	-0.0193	0.87945	0.96574
modelTM4/REAM:	5.0584	7.514	0.00889	0.017007
modelTM4/REAM:*	4.5684	8.2597	0.84129	0.5549

Even vs Odd	f-test	t-test
knmiTM4:	1.5557	0.20572
knmiREAM:	1.146	0.27781
nasaTM4:	0.8381	0.42869
nasaREAM:	0.4904	0.6605
modelTM4:	0.4455	0.5549
modelREAM:	0.0105	0.84129

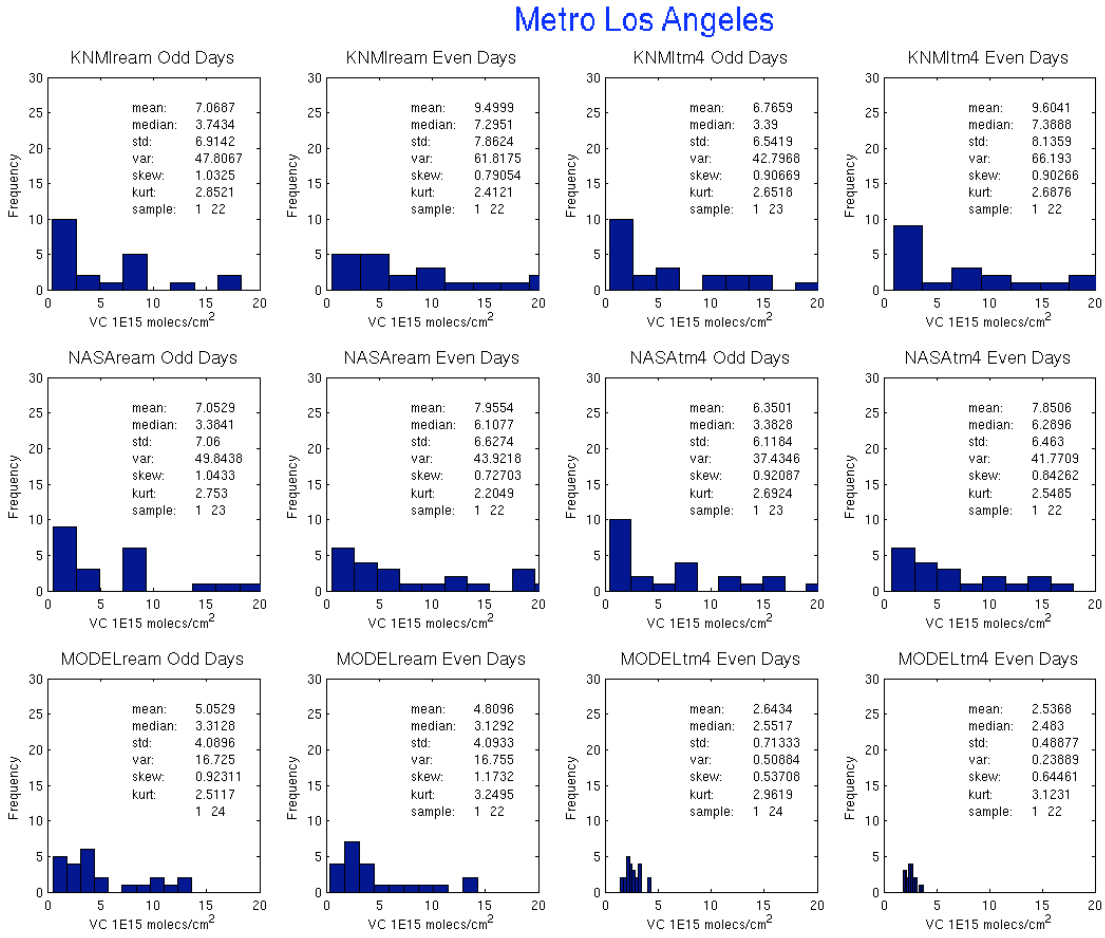


Figure A.9: Histogram for Los Angeles starting on July 7th going for 8 days

Table A.10: Statistics for inner Los Angeles starting on July 7th going for 8 days

	f-test odd days		f-test even days		t-test odd days		t-test even days	
Model vs Retrieval	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM	mTM4	mREAM
model-knmiTM4:	44.2054	0.1935	39.0287	0.7035	0.00023021	0.84	1.917e-06	0.13898
model-nasaTM4	29.3622	0.1325	84.5392	0.1556	0.00045459	0.54509	0.00027907	0.30257
model-nasaREAM:	38.9829	0.0492	57.5459	0.1472	0.00021539	0.98443	0.00017745	0.45549
model-knmiREAM:	61.9261	0.6831	56.6914	0.165	0.00027212	0.39264	1.9642e-06	0.0043237

Retrieval vs Retrieval	f-odd	f-even	t-odd	t-even
REAMnasa/knmi:	0.5158	-0.0424	0.41189	0.032457
TM4nasa/knmi	0.4308	1.0781	0.46983	0.92279
nasaTM4/REAM:	0.2591	0.4034	0.5937	0.71769
knmiTM4/REAM:	0.3343	0.3834	0.53081	0.047259
modelTM4/REAM:	35.1331	69.5216	3.3862e-05	0.00038471
modelTM4/REAM:	35.3705	69.0617	0.94588	0.75181

Even vs Odd	f-test	t-test
knmiTM4:	1.1305	0.25966
knmiREAM:	1.0559	0.082798
nasaTM4:	0.3951	0.1495
nasaREAM:	0.2666	0.48801
modelTM4:	0.8803	0.75181
modelREAM:	-0.0506	0.94588

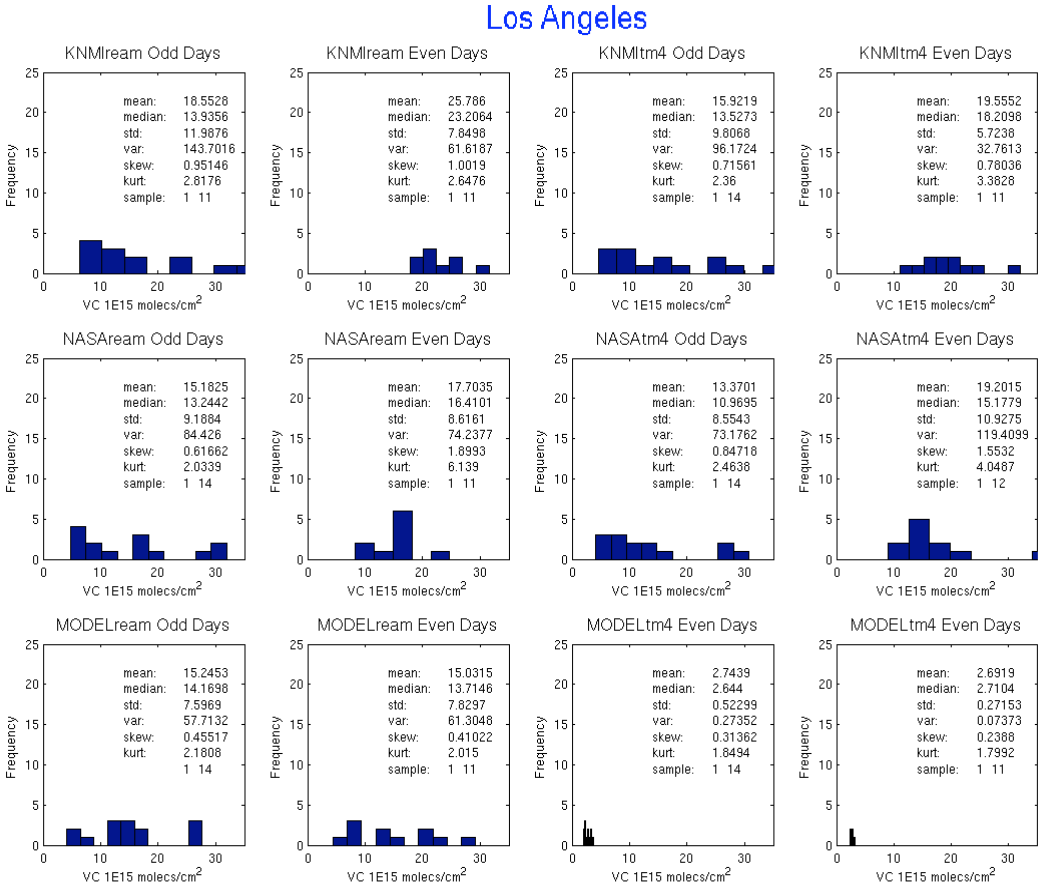


Figure A.10: Histogram for inner Los Angeles starting on July 7th going for 8 days

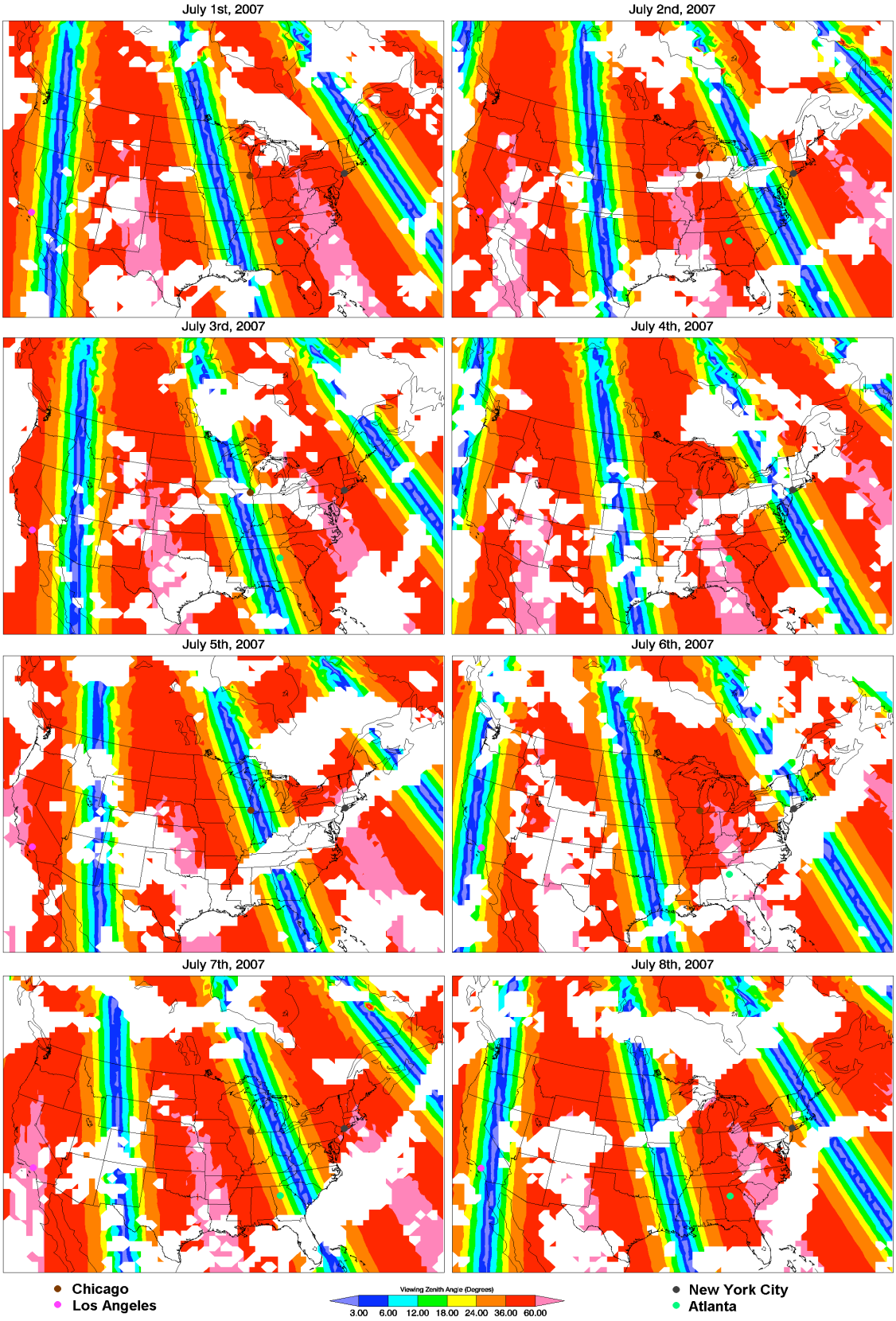

July 1st, 2007
July 2nd, 2007
July 3rd, 2007
July 4th, 2007
July 5th, 2007
July 6th, 2007
July 7th, 2007
July 8th, 2007
Chicago
Los Angeles
Viewing Zenith Angle (Degrees)
3.00 6.00 12.00 18.00 24.00 36.00 60.00
New York City
Atlanta

Figure A.11: Viewing Zenith Angle of OMI from July 1st to July 8th

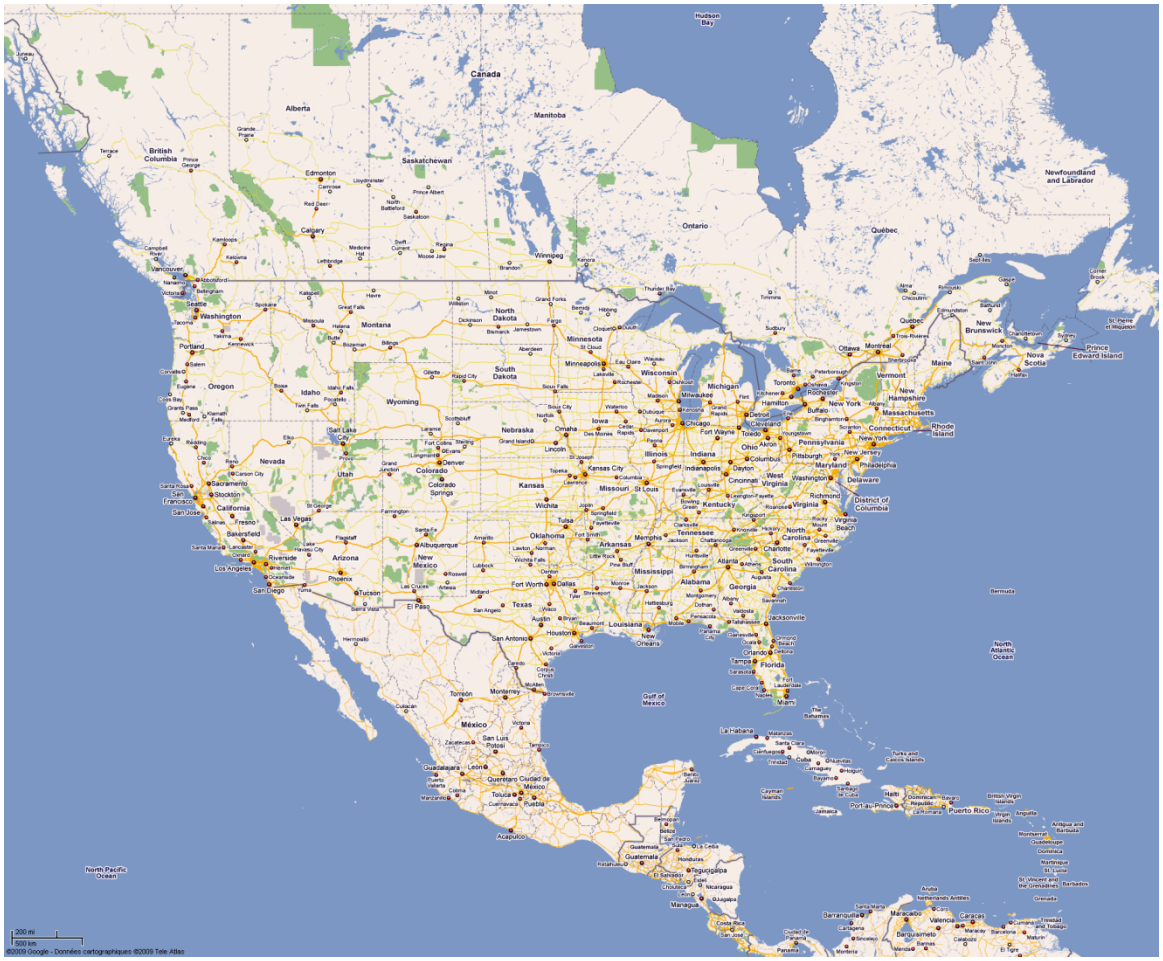


Figure A.12: Google Map of North America.

REFERENCES

Boersma, K. F., H. J. Eskes, and E. J. Brinksma (2004), Error analysis for tropospheric NO_2 retrieval from space, *J. Geophys. Res., 109*, D04311, doi:10.1029/2003JD003962.

Boersma, K. F., D. J. Jacob, H. J. Eskes, R. W. Pinder, J. Wang, and R. J. van der A (2007), Intercomparison of SCIAMACHY and OMI tropospheric NO_2 columns: Observing the diurnal evolution of chemistry and emissions from space, *J. Geophys. Res., 113*, D16S26, doi:10.1029/2007JD008816.

Boersma, K. F., H. J. Eskes, J. P. Veefkind, E. J. Brinksma, R. J. van der A, M. Sneep, G. H. J. van den Oord, P. F. Levelt, P. Stammes1, J. F. Gleason, and E. J. Bucsela, (2008). Near-real time retrieval of tropospheric NO_2 from OMI, *Atmos. Chem. Phys., 7*, 2103–2118

Boersma, K. F., D. J. Jacob, M. Trainic, Y. Rudich, I. DeSmedt, R. Dirksen, and H. J. Eskes (2009), Validation of urban NO_2 concentrations and their diurnal and seasonal variations observed from space (SCIAMACHY and OMI sensors) using in situ measurements in Israeli cities, *Atmos. Chem. Phys. Discuss., 9*, 4301–4333

Boersma K. Folkert, Ellen Brinksma, Eric J.Bucsela, Edward A. Celarier, James F. Gleason, Pieternel Levelt, J. Pepijn Veefkind, Mark O. Wenig (2008), OMI README file, *http://toms.gsfc.nasa.gov/omi/no2/OMNO2_readme.pdf*

Boersma, K. F., Bucsela, E. J., Brinksma, E. J., and Gleason, J. F. (2002): NO2, in: OMI Algorithm Theoretical Basis Document, vol. 4, OMI Trace Gas Algorithms, ATB-OMI-04, Version 2.0, edited by: K. Chance, 13–36, *NASA Distrib. Active Archive Cent.,* Greenbelt, Md., Aug.

Bogumil, K., J. Orphal, S. Voigt, H. Bovensmann, O.C. Fleischmann, M. Hartmann, T. Homann, P. Spietz and J.P. Burrows (1999), Reference Spectra of Atmospheric Trace gases measured with the SCIAMACHY PFM satellite spectrometer, *Proc. Europ. Sympos. Atm. Meas. Space, ESA-WPP-161 Vol. II*, 443-446

Bucsela, E. J., E. A. Celarier, M. O. Wenig, J. F. Gleason, J. P. Veefkind, K. F. Boersma, and E. J. Brinksma (2006), Algorithm for NO_2 vertical column retrieval from the Ozone Monitoring Instrument*, IEEE Trans. Geo. Rem. Sens., 44(5)*, 1245-1258, Special Issue on the EOS Aura Mission.

Bucsela, E. J., A. E. Perring, R. C. Cohen, K. F. Boersma, E. A. Celarier, J. F. Gleason, M. O. Wenig, T. H. Bertram, P. J. Wooldridge, R. Dirksen, and J. P. Veefkind (2008), Comparison of tropospheric NO_2 from in-situ aircraft measurements with near-real time and standard product data from OMI, *J. Geophys. Res., 113*, D16S31, doi:10.1029/2007JD008838.

Celarier E. A., E. J. Brinksma, J. F. Gleason, J. P. Veefkind (2008), Validation of Ozone Monitoring Instrument nitrogen dioxide columns, *J. Geophys. Res., VOL. 113*, D15S15, doi:10.1029/2007JD008908

Chance, K., and R.J.D. Spurr (1997), Ring Effect Studies: Rayleigh Scattering, Including Molecular Parameters for Rotational Raman Scattering, and the Fraunhofer Spectrum, *Appl. Opt. 36*, 5224-5230

Choi et al.(2005). Evidence of lightning NOx and convective transport of pollutants in satellite observations over North America, *Geophys. Res. Lett.,32*, L02805, doi:10.1029/2004GL021436

Harder, J.W. and J.W. Brault,(1997) Atmospheric measurements of water vapor in the 442-nm region, *J. Geophys. Res. 102*, 6245 – 6252

Ionov, D. V., Y. M. Timofeyev, V. P. Sinyakov, V. K. Semenov, F. Goutail, J.-P. Pommereau, E. J. Bucsela, E. A. Celarier, and M. Kroon (2008), Ground-based validation of EOS-Aura OMI NO_2 vertical column data in the midlatitude mountain ranges of Tien Shan (Kyrgyzstan) and Alps (France), *J. Geophys. Res., 113*, D15S08, doi:10.1029/2007JD008659.

Martin, R. V., D. J. Jacob, K. Chance, T. P. Kurosu, P. I. Palmer, and M. J. Evans (2003), Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns, *J. Geophys. Res., 108(D17)*, 4537, doi:10.1029/2003JD003453.

Napelenok S. L., R. W. Pinder, A. B. Gilliland, and R. V. Martin (2008), A method for evaluating spatially-resolved NOx emissions using Kalman filter inversion, direct sensitivities, and space-based NO_2 Observations, *Atmos. Chem. Phys., 8*, 5603–5614

Newnham D.A. and J. Ballard (1998), Visible absorption cross-sections and integrated absorption intensities of molecular oxygen O2 and O4, J. Geophys. Res., 103, 28801-28816

Stavrakou, T., J.-F. Muller, K. F. Boersma, I. De Smedt, and R. J. van der A (2008), Assessing the distribution and growth rates of NOx emission sources by inverting a 10-year record of NO_2 satellite columns, *Geophys. Res. Lett.*, VOL. 35, L10801, doi:10.1029/2008GL033521

Wang, Y., M. B. McElroy, R. V. Martin, D. G. Streets, Q. Zhang, and T.-M. Fu (2007), Seasonal variability of NOx emissions over east China constrained by satellite observations: Implications for combustion and microbial sources, *J. Geophys. Res., 112*, D06301, doi:10.1029/2006JD007538.

Wenig, M., S. Ku¨hl, S. Beirle, E. Bucsela, B. Ja¨hne, U. Platt, J. Gleason, and T. Wagner (2004), Retrieval and analysis of stratospheric NO_2 from the Global Ozone Monitoring Experiment, *J. Geophys. Res., 109*, D04315, doi:10.1029/2003JD003652.

Vandaele A.C., C. Hermans, P.C. Simon, M. Carleer, R. Colin, S. Fally, M.F. Mérienne, A. Jenouvrier, and B. Coquart (1998), Measurements of the NO2 absorption cross-section from 42000 cm-1 to 10000 cm-1 (238-1000 nm) at 220 K and 294 K, *J. Quant. Spectrosc. Radiat. Transfer 59*, 171-184

Van der A, R. J., H. J. Eskes, K. F. Boersma, T. P. C. van Noije, M. Van Roozendael, I. De Smedt, D. H. M. U. Peters, and E. W. Meijer (2008), Trends, seasonal variability and dominant NOx source derived from a ten year record of NO_2 measured from space, J. *Geophys. Res., 113*, D04302, doi:10.1029/2007JD009021.

Zhao, C., Y. Wang, Y. Choi, and T. Zeng (2009), Impact of convective transport and lightning NOx production over North America: dependence on cumulus parameterizations, *Atmos. Chem. Phys. Discuss., 9*, 2289–2317

Zhao, C., and Y. Wang (2009), Assimilated inversion of NOx emissions over east Asia using OMI NO2 column measurements, *Geophys. Res. Lett., 36*, L06805, doi:10.1029/2008GL037123.